

Oracle Adapter Guide

Informatica® PowerExchange®
(Version 8.6)

Copyright (c) 1998–2008 Informatica Corporation. All rights reserved.

This software and documentation contain proprietary information of Informatica Corporation and are provided under a license agreement containing restrictions on use and disclosure and are also protected by copyright law. Reverse engineering of the software is prohibited. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica Corporation. This Software may be protected by U.S. and/or international Patents and other Patents Pending.

Use, duplication, or disclosure of the Software by the U.S. Government is subject to the restrictions set forth in the applicable software license agreement and as provided in DFARS 227.7202-1(a) and 227.7702-3(a) (1995), DFARS 252.227-7013(c)(1)(ii) (OCT 1988), FAR 12.212(a) (1995), FAR 52.227-19, or FAR 52.227-14 (ALT III), as applicable.

The information in this product or documentation is subject to change without notice. If you find any problems in this product or documentation, please report them to us in writing.

Informatica, PowerCenter, PowerCenterRT, PowerCenter Connect, PowerCenter Data Analyzer, PowerExchange, PowerMart, Metadata Manager, Informatica Data Quality, Informatica Data Explorer, Informatica B2B Data Exchange and Informatica On Demand are trademarks or registered trademarks of Informatica Corporation in the United States and in jurisdictions throughout the world. All other company and product names may be trade names or trademarks of their respective owners.

Copyright, Byte Designs Ltd. All rights reserved.

This product includes ICU software which is copyright (c) 1995-2003 International Business Machines Corporation and others. All rights reserved. Permissions and limitations regarding this software are subject to terms available at <http://www-306.ibm.com/software/globalization/icu/license.jsp>.

The product includes the zlib library copyright (c) 1995-2005 Jean-loup Gailly and Mark Adler.

DISCLAIMER: Informatica Corporation provides this documentation “as is” without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of non-infringement, merchantability, or use for a particular purpose. Informatica Corporation does not warrant that this software or documentation is error free. The information provided in this software or documentation may include technical inaccuracies or typographical errors. The information in this software and documentation is subject to change at any time without notice.

Table of Contents

Preface	vii
Informatica Resources	vii
Informatica Customer Portal	vii
Informatica Documentation	vii
Informatica Web Site	vii
Informatica Knowledge Base	viii
Informatica Global Customer Support	viii
 Chapter 1: Overview	 1
PowerExchange Batch	1
Change Data Capture with PowerExchange Change and Real-time	1
Phases of Change Propagation	2
PowerExchange Oracle CDC Components	2
PowerExchange Changed Data Capture Data-Flow Components	3
 Chapter 2: Bulk Movement of Data	 5
Accessing Oracle	5
Testing a Remote PowerExchange Listener	5
Sourcing Oracle Data	5
Creating Personal Metadata	6
 Chapter 3: Configuring Oracle Change Data Capture	 7
Supported Operating Systems	7
Configuring Oracle for Oracle CDC	7
Restrictions and Prerequisites	8
Oracle ARCHIVELOG Mode	8
Oracle Transaction_auditing Parameter	8
Required Oracle Privileges	9
Supplemental Log Data	9
Configuring LogMiner	10
Copying the Oracle Catalog	10
Configuring PowerExchange for Oracle CDC	10
Configuring Oracle CDC Real-time Extraction Mode	11
Configuring Oracle CDC Batch Extraction Mode	11
Configuring Oracle CDC Continuous Extraction Mode	11
Configuration File Statements	12
Migrating to Condense File Extraction Modes	20
Oracle Datatypes	22
SQL Loader Restrictions	23
Performance Considerations	23

Chapter 4: Registering CDC Sources	25
Introduction	25
Registering Sources for Change Data Capture	25
Oracle Registration Process	26
Registering Source Data	26
Adding a Capture Registration to an Existing Group	28
Deleting a Registration Group	28
Deleting a Registration Entry	29
Viewing a Registration Group	29
Viewing a Capture Registration	29
Extraction Maps	29
Defining the Extraction Process	30
Adding an Extraction Group	30
Specify Extraction Criteria	32
Adding an Extraction Definition to an Existing Group	32
Deleting an Extraction Group	32
Deleting an Extraction Definition	32
Viewing an Extraction Definition	33
Data Delivery Format	34
Creating Restart Points with DTLUAPPL	35
Running DTLUAPPL in Windows	35
Running DTLUAPPL on Linux and UNIX	35
Adding a New Application	36
Modifying an Existing Application	37
Printing an Application	37
Application Group	38
Reset to New Start Point	38
 Chapter 5: Amending CDC Definitions	 39
Overview	39
PowerExchange View of the Captured Data	39
Removing Columns from the Extraction Map	39
Changing the Underlying Data Source	40
Managing Changes to the Data Structure	40
Changing Capture Definition Properties	40
Changing Extraction Definition Properties	41
 Chapter 6: PowerExchange Condense	 43
Introduction	43
Registering for Condense	43
Migration Consideration	44
Configuring the PowerExchange Condense Task	44
PowerExchange Condense Operational Modes	44
PowerExchange Condense Output Files	45
Configuring PowerExchange Condense Parameters	46

Configuring Flexible Condense	51
Running Flexible Condense	51
Condense Group Definition File	51
Example 1	52
Example 2	53
Output Files	53
Starting and Stopping Condense	53
Starting PowerExchange Condense	53
Stopping PowerExchange Condense	54
Chapter 7: Accessing CDC Data	55
Overview	55
Oracle CDC Real-time Extraction Mode	55
Testing with the Navigator	55
Using Real-time Extraction Mode in PowerCenter	56
PowerExchange Listener Messages During Data Extract	57
Oracle CDC Condense File Extraction	57
Testing with the Navigator	57
Using Batch Extraction Mode in PowerCenter	59
Using Continuous Extraction Mode in PowerCenter	59
Chapter 8: Stopping Change Capture	61
Stopping Change Capture	61
Index	63

Preface

This manual applies to the following PowerExchange product:

- ♦ PowerExchange for Oracle®

This manual provides the information that is required to develop an Oracle bulk and data capture environment. Use it only after successfully completing all of the required steps from the *PowerExchange Installation Guide*. At this point, the PowerExchange Listener should be active on the target platform, and you should have tested the Listener by using the Ping utility.

In this manual, the steps that are required depend on what functions of the product you will use. Two strategies are available for refreshing data in a target database:

- ♦ Bulk update by using PowerCenter
- ♦ Frequent or near real-time updating by using data capture techniques

PowerCenter or other third-party software can be used to move the data to the target.

Informatica Resources

Informatica Customer Portal

As an Informatica customer, you can access the Informatica Customer Portal site at <http://my.informatica.com>. The site contains product information, user group information, newsletters, access to the Informatica customer support case management system (ATLAS), the Informatica Knowledge Base, Informatica Documentation Center, and access to the Informatica user community.

Informatica Documentation

The Informatica Documentation team takes every effort to create accurate, usable documentation. If you have questions, comments, or ideas about this documentation, contact the Informatica Documentation team through email at infa_documentation@informatica.com. We will use your feedback to improve our documentation. Let us know if we can contact you regarding your comments.

Informatica Web Site

You can access the Informatica corporate web site at <http://www.informatica.com>. The site contains information about Informatica, its background, upcoming events, and sales offices. You will also find product and partner information. The services area of the site includes important information about technical support, training and education, and implementation services.

Informatica Knowledge Base

As an Informatica customer, you can access the Informatica Knowledge Base at <http://my.informatica.com>. Use the Knowledge Base to search for documented solutions to known technical issues about Informatica products. You can also find answers to frequently asked questions, technical white papers, and technical tips.

Informatica Global Customer Support

There are many ways to access Informatica Global Customer Support. You can contact a Customer Support Center through telephone, email, or the WebSupport Service.

Use the following email addresses to contact Informatica Global Customer Support:

- ♦ support@informatica.com for technical inquiries
- ♦ support_admin@informatica.com for general customer service requests

WebSupport requires a user name and password. You can request a user name and password at <http://my.informatica.com>.

Use the following telephone numbers to contact Informatica Global Customer Support:

North America / South America	Europe / Middle East / Africa	Asia / Australia
Informatica Corporation Headquarters 100 Cardinal Way Redwood City, California 94063 United States Toll Free +1 877 463 2435 Standard Rate Brazil: +55 11 3523 7761 Mexico: +52 55 1168 9763 United States: +1 650 385 5800	Informatica Software Ltd. 6 Waltham Park Waltham Road, White Waltham Maidenhead, Berkshire SL6 3TN United Kingdom Toll Free 00 800 4632 4357 Standard Rate Belgium: +32 15 281 702 France: +33 1 41 38 92 26 Germany: +49 1805 702 702 Netherlands: +31 306 022 797 United Kingdom: +44 1628 511 445	Informatica Business Solutions Pvt. Ltd. Diamond District Tower B, 3rd Floor 150 Airport Road Bangalore 560 008 India Toll Free Australia: 1 800 151 830 Singapore: 001 800 4632 4357 Standard Rate India: +91 80 4112 5738

CHAPTER 1

Overview

This chapter includes the following topics:

- ♦ PowerExchange Batch, 1
- ♦ Change Data Capture with PowerExchange Change and Real-time, 1
- ♦ Phases of Change Propagation, 2

PowerExchange Batch

PowerExchange Batch allows you to perform bulk moves of Oracle data from data source to target. You can move data from Linux, UNIX, and Windows systems and move data to Linux, UNIX, Windows, z/OS, and i5/OS (AS/400) systems.

Change Data Capture with PowerExchange Change and Real-time

The logical extensions of the PowerExchange Batch process are PowerExchange Change, to retrieve updates at user-defined intervals, and PowerExchange Real-time, the near real-time version.

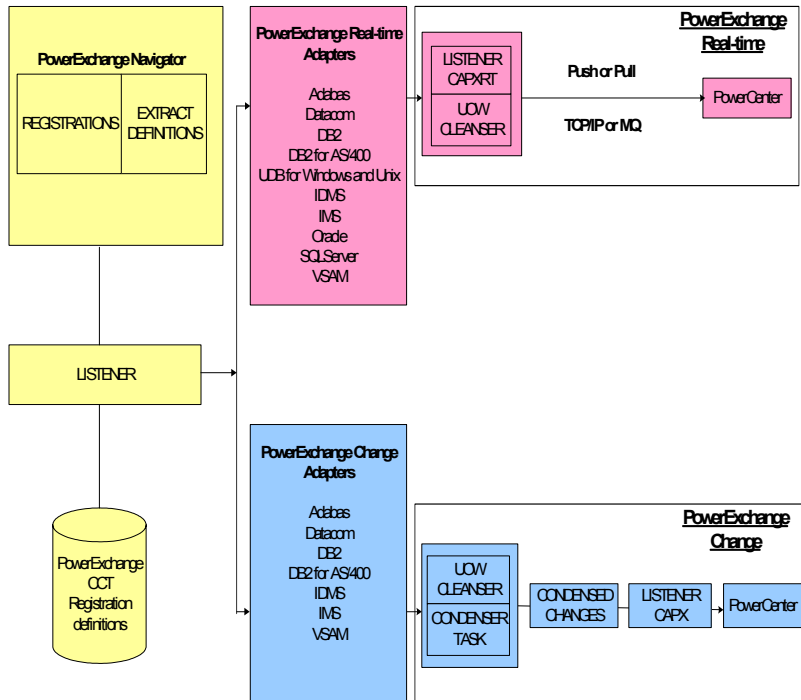
Repeating the load process through PowerExchange Batch can be costly in terms of space, time and human resources. Capturing additions, amendments, and deletions to source data and data structures can mitigate this problem. When capturing changes, verify that if changes have taken place, only the latest change is applied.

A number of rules are required to clearly identify changes and their start points. The change management process impacts operational systems and procedures. For continuous identification and capture of changes, this process must be well integrated.

A variety of components work together to propagate changed data from a source to a target:

- ♦ PowerExchange Change uses the CAPX access method to read changed data from condensed files in batch extraction mode. PowerExchange only reads closed condense files in batch extraction mode.
- ♦ PowerExchange Change also uses the CAPXRT access to read changed data from condensed files in continuous extraction mode. PowerExchange read closed and opened condense files in continuous extraction mode allowing near real-time access to changed data.
- ♦ PowerExchange Real-time uses the CAPXRT access method to read changed data directly from the Oracle redo logs using Oracle LogMiner.

PowerExchange Change and PowerExchange Real-time make data available to be applied to target databases but do not apply the data directly. In these instances, you can apply changes with PowerCenter or a partner tool.



Phases of Change Propagation

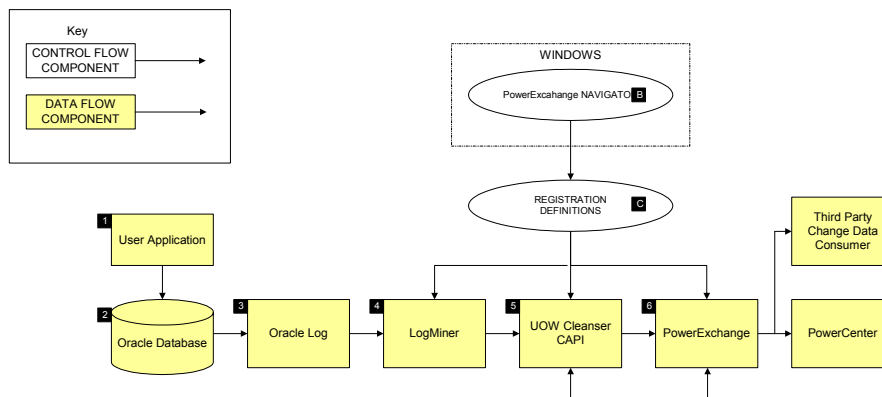
Changed Data Capture (CDC) is the process of capturing changes made to the source database and creating an inventory of those changes in an internal log.

PowerExchange Oracle CDC Components

PowerExchange Oracle CDC sources data from Oracle redo logs using Oracle LogMiner sessions.

Figure 1-1 depicts the PowerExchange Oracle Change Data Capture architecture for an Oracle source that can be applied to a target database using PowerCenter or a partner product:

Figure 1-1. Architecture of the PowerExchange Oracle CDC Components



PowerExchange Changed Data Capture Data-Flow Components

Table 1-1 describes the data flow components. These components handle the changed data as it progresses through the propagation process. Note that some of these components, such as the user application and Oracle source database, are not PowerExchange Change Capture components.

Table 1-1. Data-Flow Components Involved in Change Propagation

Component	Description
User Application	Any software that updates the source database on a continuous basis.
Oracle Database	Oracle database where the source data resides. PowerExchange Oracle CDC can capture changes from more than one source database or data file. Oracle data is captured by utilizing existing Oracle LogMiner technology.
Oracle Log	Oracle redo logs are an essential part of the PowerExchange Oracle CDC capture process. Archive Logging is set up at an early stage of the installation process.
LogMiner	Oracle LogMiner is a required product for PowerExchange Oracle CDC. Oracle Redo logs are made available to PowerExchange software which will extract the data from specified points in time delimited by Oracle System Change Numbers.
UOW Cleanser CAPI	Unit of work cleanser ensures that units of work are processed in the sequence they are committed, and that rolled back transactions are discarded. The Oracle Consumer API (CAPI) connects to the PowerExchange capture registration definitions of the tables and columns of interest, then connects to Oracle and Log Miner and uses the restart information to determine which records should be returned.
PowerExchange	PowerExchange can direct data returned to a number of targets. These can involve passing data to a partner tool to apply the changes to the target or using PowerCenter to apply the data changes directly to the target.
PowerExchange Navigator	The PowerExchange Navigator is used to specify the PowerExchange registration, locations, users, and user groups.
Registration Definitions	Central repository holds information on all the data sources and locations. The repository can be created on whatever platform is preferable.

The following document describes the necessary steps for developing a data capture environment:

- ◆ dbmover.cfg configuration parameters
- ◆ Registering a data source
- ◆ Oracle prerequisites
- ◆ Configuring Oracle Log Miner
- ◆ Materialization of tables using the Bulk Move facility
- ◆ Registering a Restart Token
- ◆ Data access

For more information, see “Configuring Oracle Change Data Capture” on page 7.

CHAPTER 2

Bulk Movement of Data

This chapter includes the following topics:

- ♦ Accessing Oracle, 5
- ♦ Testing a Remote PowerExchange Listener, 5
- ♦ Sourcing Oracle Data, 5
- ♦ Creating Personal Metadata, 6

Accessing Oracle

The following information must be available before you can source Oracle data:

- ♦ Oracle SID
- ♦ Database Service Name
- ♦ Tables to be sourced and whether they are grouped by a particular owner
- ♦ User ID and password to access these tables

Testing a Remote PowerExchange Listener

To use a remote PowerExchange Listener, it must first be started. You can check if a remote listener is running and your connectivity to the PowerExchange Listener by entering the following command at the command line:

```
dtlrexe loc=node prog=ping
```

Replace *node* with the node name in the dbmover.cfg file that points to the remote platform.

Sourcing Oracle Data

Complete the following tasks to use Oracle data as a source:

- ♦ Get the valid user ID and password for the source database.
- ♦ Get any database specific information, such as the Oracle SID.

- ◆ Ping the IP address of the source machine to test your connectivity.
- ◆ Add a node name for the location in dbmover.cfg.
- ◆ Use the command “dtlrex loc=*node* prog=ping uid=*userID* pwd=*pwd*” to test that the remote PowerExchange Listener is started, where *node* is the NODE name from the dbmover.cfg file.
- ◆ If necessary, create the Oracle LogMiner environment.
- ◆ After checking that metadata returns successfully, use the PowerExchange Navigator to create a personal metadata profile for the Oracle source database and view data. For more information, see “Creating Personal Metadata” on page 6.
- ◆ Use PowerCenter to execute the request to move data.

Creating Personal Metadata

To create personal metadata:

1. In the Resource Explorer pane, right-click **Personal Metadata** and click **Add Personal Metadata**.
2. In the Personal Metadata - Name dialog box, enter the appropriate details for the data that you want to source and click **Next**.

Field	Description
Name	Name used locally on the machine to identify the location information for this group. This should be as meaningful as possible.
Location	Location of the PowerExchange Listener on which the Oracle instance resides. This will be defined in dbmover.cfg.
Type	Select Oracle.
UserID/Password	Enter a user ID and password if required by the security setting in the dbmover configuration file. If you do not enter a password when one is required, a password dialog box appears. Leaving the password blank prevents the password being stored by PowerExchange. This can cause password violations if the Personal Metadata is read after the password for the user has changed.
SQL *Net Name	Oracle Net Name.
Description	Description of the Personal Metadata request.

3. Optionally, you can filter the results that you retrieve by entering criteria in the Personal Metadata - Filters dialog box.

Field	Description
Filter Data	Enter a schema and table name to reduce the number of database tables returned, thus saving time.
Escape Character	Defines the escape character for when the resource name contains special characters.

4. Click **Finish**.

The tables that you requested display under your new Personal Metadata in the Resource Explorer. Double-click a table to display table information.

To perform a Row Test to test connectivity to your selected data source, click the **Row Test** button.

5. Click **Go**.

CHAPTER 3

Configuring Oracle Change Data Capture

This chapter includes the following topics:

- ♦ Supported Operating Systems, 7
- ♦ Configuring Oracle for Oracle CDC, 7
- ♦ Configuring PowerExchange for Oracle CDC, 10
- ♦ Migrating to Condense File Extraction Modes, 20
- ♦ Oracle Datatypes, 22
- ♦ SQL Loader Restrictions, 23
- ♦ Performance Considerations, 23

Supported Operating Systems

Oracle CDC runs on all versions of the following platforms that PowerExchange supports:

- ♦ Windows 2000
- ♦ Windows XP
- ♦ AIX
- ♦ Solaris
- ♦ HP-UX
- ♦ Linux

Configuring Oracle for Oracle CDC

To use PowerExchange for Oracle CDC, you must configure certain components of Oracle such as LogMiner. PowerExchange uses LogMiner to extract changes from the Oracle redo logs.

PowerExchange provides scripts to help configure Oracle so you can run Oracle CDC. The scripts contain instructions and considerations for configuring Oracle. Read the instructions before running any of the supplied SQL statements or changing any Oracle configuration parameters.

PowerExchange provides the following scripts in the PowerExchange install directory:

Table 3-1. Oracle CDC Sample Scripts

File Name	Description
oracapt.sql	Use for Oracle CDC when running in a non-RAC environment.
oracapt_rac.sql	Use for Oracle CDC when running in an Oracle 10g or higher RAC environment.

Note: Oracle CDC supports Oracle Real Application Cluster (RAC) environments when running Oracle Version 10.1 or higher.

Restrictions and Prerequisites

Review the following restrictions and prerequisites before using Oracle CDC:

- ♦ Use a version of Oracle, with the required fixes, supported by PowerExchange.
- ♦ Enable Oracle archive logging. For more information about Oracle archive logging, see “Oracle ARCHIVELOG Mode” on page 8.
- ♦ Enable Oracle Supplemental Logging. See “Supplemental Log Data” on page 9.
- ♦ Configure Oracle LogMiner. LogMiner internally opens a number of cursors to handle its processing. Configuring LogMiner for the first time can result in “number of open cursors exceeded” messages. You may need to increase the maximum number of open cursors within Oracle to handle the extra LogMiner processing. For more information about configuring LogMiner, see “Configuring LogMiner” on page 10.
- ♦ Oracle CDC cannot process copies of archive logs. Oracle CDC uses LogMiner continuous mining. LogMiner continuous mining reads the archive logs from the original directory to which they are written.
- ♦ The Oracle catalog is copied to the Oracle Redo Log. See “Copying the Oracle Catalog” on page 10.
- ♦ If Oracle CDC is not installed on the same machine as the Oracle instance, configure a TNS entry on the client machine with SERVER=DEDICATED in the CONNECT_DATA section of the connect descriptor. This is also required if the network is configured for Multi-Threaded Server (MTS) mode.
- ♦ PowerExchange requires the Oracle Client binaries. Oracle installs the Client binaries by default in a normal Oracle instance installation. When using the SQL*Net connectivity on a machine that does not have an Oracle instance installed, install the Oracle Client.

Oracle ARCHIVELOG Mode

Oracle must be running in ARCHIVELOG mode before you use Oracle CDC. If Oracle is not running in Archivelog mode, configure ARCHIVELOG mode. PowerExchange provides the necessary information on how to do this. For the names of the files containing this information, see Table 3-1 on page 8.

Oracle Transaction_auditing Parameter

Oracle versions 9.2.0.3 and 9.2.0.4 require the init.ora transaction_auditing parameter be set to True for Oracle CDC to work correctly. True is the default value. Oracle has opened a bug report, 3456259, to address the problem. When installing PowerExchange for Oracle CDC, check the status of this issue. If Oracle has provided a patch, that fix can be installed instead of changing init.ora or spfile.

If necessary, change init.ora to specify this value. Alternatively, if you are using a server parameter file (spfile), you can execute the ALTER SYSTEM SET statement below in an SQL*Plus session with a user ID with the required authority:

```
ALTER SYSTEM SET transaction_auditing = TRUE SCOPE=SPFILE;
```


Required Oracle Privileges

Users extracting changes from Oracle redo logs using Oracle CDC require specific Oracle system and object privileges. Create a new user specifically for Oracle CDC usage, use an existing user with the required authority, or grant the required privileges to an existing user. PowerExchange provides scripts containing the necessary SQL GRANT statements. For the file names, see Table 3-1 on page 8. Modify these scripts as necessary for your environment.

You must grant the following system privileges to any Oracle CDC user:

Table 3-2. Minimum Required Oracle System Privileges

System Privilege	Oracle Release	Description
SELECT ANY TRANSACTION	10g and higher	Required for all users extracting Oracle CDC data in real time and for all Condense tasks.
ALTER ANY TABLE	all	Required for users that, when defining capture registrations, execute the DDL generated to create the supplemental log group for the table being registered.
CONNECT	all	Required for all users extracting Oracle CDC data in real time and for all Condense tasks.
LOCK ANY TABLE	all	Required if GENRLOCK=Y is specified on the Oracle CAPI_CONNECTION statement in dbmover.cfg.

You must grant the following object privileges to any Oracle CDC user:

Table 3-3. Minimum Required Object Privileges

Object Privilege	Object Name
SELECT	Any table registered for Oracle capture. These GRANTS are not required if you specify GENRLOCK=Y on the CAPI_CONNECTION statement in dbmover.cfg.
SELECT	PUBLIC.V\$ARCHIVED_LOG
SELECT	PUBLIC.V\$DATABASE
SELECT	PUBLIC.V\$INSTANCE
SELECT	PUBLIC.V\$LOGMNR_CONTENTS
SELECT	PUBLIC.V\$NLS_PARAMETERS
SELECT	PUBLIC.V\$PARAMETER
SELECT	PUBLIC.V\$TRANSACTION
SELECT	SYS.DBA_LOG_GROUPS
SELECT	SYS.DBA_LOG_GROUP_COLUMNS
EXECUTE	SYS.DBMS_FLASHBACK
EXECUTE	SYS.DBMS_LOGMNR
EXECUTE	SYS.DBMS_LOGMNR_D

Supplemental Log Data

Supplemental logging is required for Oracle CDC. Oracle limits data capture at the table level, rather than the database level, when you use supplemental logging by table.

Only define the columns that need to be captured in the Oracle supplemental log group. When you register a table for Oracle CDC in the Navigator, PowerExchange automatically generates DDL to define the

supplemental log group. PowerExchange only includes the columns you select when registering the table in the DDL for the supplemental log group. You can execute the DDL immediately or execute it later.

Configuring LogMiner

PowerExchange uses Oracle LogMiner for Oracle CDC. Oracle suggests that you create a tablespace for the exclusive use of LogMiner. Creating a dedicated LogMiner tablespace helps prevent potential filling of the SYSTEM tablespace, a situation that can cause major service issues.

PowerExchange provides the DDL to create this in files supplied in the install. For the file names, see Table 3-1 on page 8. Read the comments in the file carefully before running any LogMiner statements.

Copying the Oracle Catalog

Oracle CDC requires a copy of the Oracle catalog on the Oracle log before you extract data. Execute the following command:

```
begin
  SYS.DBMS_LOGMNR_D.BUILD(
    options => sys.dbms_logmnr_d.store_in_redo_logs);
end;
/
```

Configuring PowerExchange for Oracle CDC

Using the PowerExchange Navigator, you create capture registrations to capture changes for Oracle tables. The capture registration process generates ORACLE SQL that you execute alter the table to add the supplemental log information for the table. Oracle uses the supplemental log information to log the change data to its logs. Oracle CDC uses Oracle LogMiner to extract captured change data from Oracle redo logs. Oracle CDC provides the following extraction options:

- ♦ **Real-time extraction mode.** When you use real-time extraction mode, Oracle CDC initiates a LogMiner continuous mining session to read changes from the Oracle redo logs. Oracle CDC continues reading changes from the Oracle redo logs until you stop the extraction.
Oracle CDC establishes a LogMiner session for each real-time mode extraction. Using real-time extraction provides the lowest latency for change data extraction and the highest impact on system resources.
- ♦ **Batch extraction mode from condense files.** When you use batch extraction mode from condense files, Oracle CDC reads changes from condense files created by the PowerExchange Condense task. Oracle CDC reads only the closed condense files that existed when the extraction starts and ends the extraction when these condense files have been processed.
Oracle CDC establishes a LogMiner session for each Condense task which then creates the condense files. The PowerExchange Listener reads the condense files. Using batch extraction mode from condense files provides the highest latency for change data extraction while minimizing the impact on system resources.
- ♦ **Continuous extraction mode from condense files.** When you use continuous extraction mode from condense files, Oracle CDC reads changes from condense files created by the PowerExchange Condense task. Oracle CDC reads data from condense files as it is being written by the Condense task, even if the condense file is not yet closed. Oracle CDC continues reading changes from condense files until you stop the extraction.
Oracle CDC establishes a LogMiner session for each Condense task which then creates the condense files. The PowerExchange Listener reads the condense files. Using continuous extraction mode from condense files provides near real-time access to change data while minimizing the impact on system resources.

For more information about the PowerExchange Condense task, see “PowerExchange Condense” on page 43.

You can configure PowerExchange for all three extraction modes providing flexibility when choosing an extraction mode for a specific application or set of tables.

Configuring Oracle CDC Real-time Extraction Mode

Complete the following steps to configure PowerExchange Oracle CDC real-time extraction mode:

1. Configure Oracle. See “Configuring Oracle for Oracle CDC” on page 7.
2. Configure PowerExchange:
 - ♦ See “CAPT_PATH and CAPT_XTRA Statements” on page 13.
 - ♦ See “ORACLEID Statement” on page 13.
 - ♦ See “CAPI_SRC_DFLT Statement” on page 14.
 - ♦ See “Required CAPI_CONNECTION Statements” on page 14.
3. Register tables for Oracle CDC. See “Oracle Registration Process” on page 26.
4. Test the configuration. See “Oracle CDC Real-time Extraction Mode” on page 55.
5. Configure the PowerCenter extraction by creating mappings and workflows. See the *PowerExchange Interfaces for PowerCenter*.
6. Establish a starting point in the change stream for the extraction. See “Creating Restart Points with DTLUAPPL” on page 35.

Configuring Oracle CDC Batch Extraction Mode

Complete the following steps to configure PowerExchange Oracle CDC batch extraction mode for condense files:

1. Configure Oracle. See “Configuring Oracle for Oracle CDC” on page 7.
2. Configure PowerExchange:
 - ♦ See “CAPT_PATH and CAPT_XTRA Statements” on page 13.
 - ♦ See “ORACLEID Statement” on page 13.
 - ♦ See “CAPI_SRC_DFLT Statement” on page 14.
 - ♦ See “Required CAPI_CONNECTION Statements” on page 14.
 - ♦ See “Configuring the PowerExchange Condense Task” on page 44.
3. Register tables for Oracle CDC selecting Part in the Condense option. See “Oracle Registration Process” on page 26. For existing registrations, see “Changing Capture Definition Properties” on page 40.
4. Start the PowerExchange Condense task. See “Starting and Stopping Condense” on page 53.
5. Test the configuration. See “Oracle CDC Condense File Extraction” on page 57.
6. If migrating from real-time extraction mode, see “Migrating to Condense File Extraction Modes” on page 20.
7. Configure the PowerCenter extraction by creating mappings and workflows. See the *PowerExchange Interfaces for PowerCenter*.
8. Establish a starting point in the change stream for the extraction. See “Creating Restart Points with DTLUAPPL” on page 35.

Configuring Oracle CDC Continuous Extraction Mode

Complete the following steps to configure PowerExchange Oracle CDC continuous extraction mode:

1. Configure Oracle. See “Configuring Oracle for Oracle CDC” on page 7.
2. Configure PowerExchange:
 - ♦ See “CAPT_PATH and CAPT_XTRA Statements” on page 13.
 - ♦ See “ORACLEID Statement” on page 13.
 - ♦ See “CAPI_SRC_DFLT Statement” on page 14.
 - ♦ See “Required CAPI_CONNECTION Statements” on page 14.
 - ♦ See “CAPI_CONNECTION Statement for Continuous Extraction Mode” on page 19.
 - ♦ See “Configuring the PowerExchange Condense Task” on page 44.
3. Register tables for Oracle CDC selecting Part in the Condense option. See “Oracle Registration Process” on page 26. For existing registrations, see “Changing Capture Definition Properties” on page 40.
4. Start the PowerExchange Condense task. See “Starting and Stopping Condense” on page 53.
5. Test the configuration. See “Oracle CDC Condense File Extraction” on page 57.
6. If migrating from real-time or batch extraction mode, see “Migrating to Condense File Extraction Modes” on page 20.
7. Configure the PowerCenter extraction by creating mappings and workflows. See the *PowerExchange Interfaces for PowerCenter*.
8. Establish a starting point in the change stream for the extraction. See “Creating Restart Points with DTLUAPPL” on page 35.

Configuration File Statements

PowerExchange uses the specifications provided in the configuration file, `dbmover.cfg`, to configure the operation of various PowerExchange tasks as well as their communication with other PowerExchange tasks. For a complete list of all configuration file parameters, see the *PowerExchange Reference Manual*.

Oracle CDC and the PowerExchange Exchange Listener use parameters specified in the configuration file to:

- ♦ Connect to Oracle databases.
- ♦ Determine the directory in which to store capture registrations, extraction maps, capture files, and condense files.
- ♦ Get connection information in order to extract changes from Oracle redo logs and from condense files.

The Condense task has an additional configuration file, `dtlca.cfg`. See “Configuring the PowerExchange Condense Task” on page 44.

Configure the following configuration parameters for Oracle CDC:

Table 3-4. Oracle CDC dbmover.cfg Statements

Statement	Required	Description
CAPT_PATH	No	PowerExchange stores registrations, capture files such as the CCT and CDCT, condense files, and Condense task checkpoint files in the specified directory. Default is the PowerExchange install directory. See “CAPT_PATH and CAPT_XTRA Statements” on page 13.
CAPT_XRTA	No	PowerExchange stores capture extraction maps in the specified directory. Default is the PowerExchange install directory. See “CAPT_PATH and CAPT_XTRA Statements” on page 13.
CAPI_SRC_DFLT	Yes	Specifies the default CAPI_CONNECTION PowerExchange uses when none is supplied. Using CAPI_SRC_DFLT, you can specify default CAPI_CONNECTION statements for different source types in a single configuration file. See “CAPI_SRC_DFLT Statement” on page 14.

Table 3-4. Oracle CDC dbmover.cfg Statements

Statement	Required	Description
CAPI_CONNECTION	Yes	PowerExchange requires CAPI_CONNECTION statements for real-time extraction mode, continuous extraction mode from condense files, and the Condense task. See "Required CAPI_CONNECTION Statements" on page 14 and "CAPI_CONNECTION Statement for Continuous Extraction Mode" on page 19.
ORACLEID	Yes	Specifies the name and connection information for any Oracle instances which Oracle CDC processes. See "ORACLEID Statement" on page 13.

CAPT_PATH and CAPT_XTRA Statements

By default, PowerExchange creates all capture files, registrations, and extraction maps in the PowerExchange install directory. Using the following parameters in the PowerExchange configuration file, dbmover.cfg, you can specify different directories for these objects:

Parameter	Description
CAPT_PATH	PowerExchange stores registrations, capture files such as the CCT and CDCT, condense files, and Condense task checkpoint files in the specified directory. Create this directory, if it does not exist, prior to using Oracle CDC. Default is the PowerExchange install directory.
CAPT_XTRA	PowerExchange stores capture extraction maps in the specified directory. Create this directory, if it does not exist, prior to using Oracle CDC. Default is the PowerExchange install directory.

Specify these parameters with unique directories to separate CDC information from the PowerExchange binary files. Separating these objects from the PowerExchange binaries can ease the software migration process since data objects and files are not inter-mingled with the PowerExchange code.

ORACLEID Statement

Oracle CDC requires an ORACLEID statement for each Oracle instance which is used for CDC. Add the following statement to the dbmover.cfg configuration file on the machine running Oracle CDC:

```
ORACLEID=(collection_id,oracle_sid,connect_string,capture_connect_string)
```

The following table describes the variables in the ORACLEID statement:

Variable	Description
capture_connect_string	Database Service Name, such as OEMREP.INTERNAL.INFA.COM. Enables Oracle capture code to connect to the Oracle_SID defined above. Code the <i>capture_connect_string</i> as follows: 1. When the job used to capture Oracle change data runs on a machine that is remote from the Oracle instance. 2. When the job used to capture Oracle change data runs on a machine with more than one Oracle Instance, and the instance that contains the data to be captured is <i>not</i> the default instance on the machine, that is, when the instance is <i>not</i> the one pointed to by the ORACLE_SID environment variable. In the above instances, <i>capture_connect_string</i> must be coded in the dbmover.cfg file of the job involved in retrieving the Oracle data.
collection_id	<i>Collection_id</i> that matches that requested in the capture registration process.

Variable	Description
connect_string	Database Service Name, such as OEMREP.INTERNAL.INFA.COM. Enables the Repository code to connect to the Oracle SID defined above. The <i>Connect_string</i> must be coded when the Navigator is used to create an Oracle CDC registration and <ol style="list-style-type: none"> 1. The PowerExchange Listener involved in retrieving the Oracle data is remote from the Oracle instance. 2. The PowerExchange Listener involved exists on a machine with more than one Oracle Instance, and the instance which contains the data to be captured is <i>not</i> the default instance on the machine (for example, the instance is <i>not</i> the one which is pointed to by the listener's ORACLE_SID environment variable. In the above instances, <i>connect_string</i> must be coded in the dbmover.cfg file of the listener involved in retrieving the Oracle data. In these cases, if row test is to be run to verify the registration, <i>capture_connect_string</i> must also be coded in the dbmover.cfg file of the PowerExchange Listener involved in retrieving the Oracle data.
Oracle_SID	Name of the Oracle database containing the registered tables.

Add ORACLEID statements for all Oracle instances using Oracle CDC.

Note: You can enter a maximum of 20 ORACLEID statements in the configuration file.

CAPI_SRC_DFLT Statement

You can code multiple CAPI connection statements in a single DBMOVER configuration file. Use the CAPI_SRC_DFLT to specify the default CAPI_CONNECTION statement for a specific source type. PowerExchange uses the CAPI_SRC_DFLT statement for a specific source type when no CAPI_CONNECTION override is supplied. The following example shows CAPI_SRC_DFLT statements for multiple source types in a single DBMOVER configuration file:

```
CAPI_SRC_DFLT= (MSQL, sql_name)
CAPI_SRC_DFLT= (UDB, db2_name)
CAPI_SRC_DFLT= (ORA, uowc_name)
```

To use continuous extraction mode for condense files, the extraction must use a CAPX CAPI_CONNECTION statement. If the CAPI_SRC_DFLT statement points to a source CAPI_CONNECTION, specify the CAPI connection override for the extraction.

PowerExchange Condense uses source CAPI_CONNECTION statements. If the CAPI_SRC_DFLT statement points to a CAPX CAPI_CONNECTION, specify the appropriate source CAPI_CONNECTION name in the CONN_OVR parameter in the PowerExchange Condense configuration file.

Required CAPI_CONNECTION Statements

PowerExchange uses CAPI connection statements to configure CDC processing. Oracle CDC requires the following CAPI_CONNECTION statements:

- ◆ UOWC for the UOW Cleanser
- ◆ ORCL for Oracle real-time extraction mode and for the Condense task

Add the UOWC and ORCL CAPI_CONNECTION statements to the PowerExchange configuration file, dbmover.cfg, on the machine running Oracle CDC. For example:

```
CAPI_CONNECTION= (NAME=uowc_name, TYPE= (UOWC, CAPINAME=orcl_name))
CAPI_CONNECTION= (NAME=orcl_name, TYPE= (ORCL, REPNODE=local, ORACOLL=Collection_id))
```

The orcl_name specification on the UOW Cleanser parameter CAPINAME must be the name specified in the Oracle CAPI_CONNECTION NAME parameter.

Collection_id must match the collection identifier specified in the ORACLEID statement which matches the collection identifier specified in the Registration Group containing the capture registrations.

Note: PowerExchange does not require CAPI_CONNECTION statements on the PowerExchange Navigator machine. You can register Oracle tables for CDC and perform database row test functions in the Navigator by using the PowerExchange Listener where Oracle CDC is running.

ORCL CAPI_CONNECTION Parameters

Use the ORCL CAPI_CONNECTION statement to specify parameters for Oracle CDC real-time extraction mode and for Oracle Condense.

Also specify a UOWC CAPI_CONNECTION statement. The CAPINAME parameter of the UOWC statement must specify the same value as the NAME parameter in the ORCL statement.

The ORCL statement has the following syntax:

```
CAPI_CONNECTION= (
  NAME=name,
  [DLLTRACE=trace_id] ,
  [TRACE=trace] ,
  TYPE= (ORCL,
    [ARRAYSIZE=nnnn] ,
    [BYPASSUF=Y|N] ,
    [CATBEGIN=hh:mm] ,
    [CATEND=hh:mm] ,
    [CATINT=mins] ,
    [COMMITINT=mins] ,
    [GENRLOCK=Y|N] ,
    [LOGDEST=dest] ,
    [LGTHREAD=thread] ,
    [ONLINECAT=Y|N] ,
    ORACOLL=collection_id,
    [SELRETRY=retry] ,
    [SGLINST=ora_instance]
  )
)
```

Parameter	Required	Default	Description
ARRAYSIZE	No	100	Use this parameter to control the size of the prefetch array Oracle capture uses to read the Oracle archive logs. The default directs Oracle Capture to request that Oracle prefetch 100 rows at a time. Valid values are 0 to $(2^{31}) - 1$. Specify 0 to turn prefetch off. Note: Specifying a value of less than 100 can adversely affect Oracle Capture performance. Prefetch should only be turned off at the request of Informatica Global Customer Support.
BYPASSUF	No	N	If your Oracle instance contains Global Temporary tables and any of these tables contain LOB columns that can be included in the table row (this is the default), then specify BYPASSUF=Y. Otherwise, you do not have to specify this parameter.
CATBEGIN	No	00:00	The earliest time of day at which Oracle can attempt to write the Oracle catalog to the Oracle redo log. This parameter uses the following format: HH24 :MM
CATEND	No	24:00	The latest time of day at which Oracle can attempt to write the Oracle catalog to the Oracle redo log. This parameter uses the following format: HH24:MM format CATEND must be > CATBEGIN. To accept the default of 24:00 do not specify CATEND, otherwise specify CATEND with any valid time from 00:01 to 23:59.
CATINT	No	1440	The number of minutes between attempts to write the Oracle catalog to the Oracle redo log. If CATINT indicates it is time to write the catalog to the redo log and it is outside the bounds specified between CATBEGIN and CATEND, then it will not be written. Valid values are 1-1440.

Parameter	Required	Default	Description
COMMITINT	No	5 Minutes	Specifies the number of minutes between Oracle Capture commit points. While Oracle Capture does not update Oracle resources, it still has a presence on the Oracle redo log. Without occasional commits, the start point in the redo log of the current Oracle Capture becomes the restart point of the next instantiation. This lengthens restart processing. Since Oracle Capture is using OCI to create its own environment to log into Oracle, it is separate from any other Oracle task in the process. This means that there will be no commit conflicts between Oracle Capture and other tasks in the process. Valid values are 1 to 60.
DLLTRACE	No	none	Specifies the name of the TRACE= statement that activates internal DLL tracing for this specific CAPI. Only specify this parameter when instructed to do so by Informatica Global Customer Support.
GENRLOCK	No	N	Specifies whether or not to issue a <code>LOCK TABLE IN EXCLUSIVE MODE</code> before returning a restart point in a call to <code>CAPI_GenRestartInfo</code> . Specify Y or N. Specifying N means that, if the table is in use, whether or not changes will be captured for the transaction which is currently using the table depends on the SCN value captured by <code>CAPI_GenRestartInfo</code> . If the captured SCN value is equal to that of the current transaction, then that transaction will be captured. If the captured SCN value is greater than that of the current transaction, then that transaction will not be captured. Note that there is no way to control what will happen to the current transaction if <code>GENRLOCK</code> is N. Note: Due to the security requirements necessary to issue a <code>LOCK TABLE</code> command, the default is N.
LOGDEST	No	NONE	Specify this option when you want to force Oracle Capture to use archive logs from a specific log archive destination. For example, if you want to force Oracle Capture to use archive logs from the destination set by the <code>init.ora log_archive_dest_3</code> parameter, specify <code>LOGDEST=3</code> . Note that the archive log destination you specify must be local to the instance that runs Oracle Capture. Valid values are 1 to 10.
LGTHREAD	No	NONE	Specify this option in an Oracle RAC environment when you want to force Oracle Capture to use archive logs created by a specific instance. Specify the instance number of the instance to which the archive logs belong that you wish Oracle Capture to use. Valid values are 1 to 2147483647.
NAME	Yes	none	Specifies a unique name for the <code>CAPI_CONNECTION</code> statement. Maximum is 8 characters.
ORACOLL	Yes	none	Specifies the Oracle collection identifier on the <code>ORACLEID</code> statement.

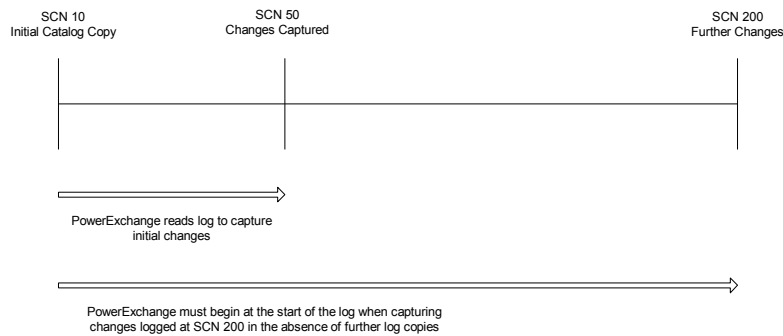
Parameter	Required	Default	Description
ONLINECAT	No	N	<p>Specifying ONLINECAT=Y causes Oracle Capture to instruct LogMiner to use the online catalog to format log data, instead of using the catalog copy on the archive log.</p> <p>The advantage of specifying ONLINECAT=Y is that Oracle Capture initializes much faster. LogMiner does not generate data in its shadow tables at startup, so there is less data for an Oracle Capture session to read.</p> <p>The disadvantage of specifying ONLINECAT=Y is that LogMiner does not allow DDL tracking. This means that LogMiner is unable to format all the log data for tables you are capturing, if either:</p> <ul style="list-style-type: none"> - DDL for those tables has occurred during capture processing. - DDL for those tables has been recorded on the archive logs you are processing. <p>If you also use the BYPASSUF=Y parameter setting, this causes Oracle Capture to skip data that it cannot format. This can cause data loss on your targets.</p>
SELRETRY	No	1000	<p>Specifies the number of times that Oracle Capture should immediately loop back to the LogMiner call before implementing a graduated-scale wait loop.</p> <p>After the specified number of LogMiner call retries, Oracle Capture starts to wait between each subsequent retry. The wait begins at one millisecond, and gradually increases to one second. When data is actually returned from the LogMiner call, the wait is reset to 0, and the process begins again for the next call to LogMiner.</p> <p>Valid values are 0 to $(2^{31}) - 1$.</p> <p>When you specify a non-zero value, Oracle Capture uses non-blocking SQL to ensure timely shutdown of the session, when requested by the user.</p> <p>When you specify 0, Oracle Capture will not use non-blocking SQL. This improves CPU consumption at the cost of ensuring timely session shutdown requests. Oracle Capture will not honor a shutdown request until log data is returned from Oracle. This only affects quiescent systems. If update activity is occurring on the Oracle instance, you will not notice a difference in shutdown behavior.</p>
SNGLINST	No	none	<p>Specify this option when you want to force Oracle Capture to use only archive logs from a specific log archive thread and destination. It allows you to run Oracle Capture in an Oracle 9i RAC environment.</p> <p>You must specify the corresponding LGTHREAD and LOGDEST parameters. Only the changes made by the instance you specify in the LGTHREAD parameter are captured.</p> <p>If you use this parameter to capture changes in an Oracle RAC environment, you are responsible for running separate capture processes to capture change made by the other RAC instances and for determining how to properly merge these changes to your target environment.</p>
TRACE	No	none	<p>Specifies the name of the TRACE= statement that activates the common CAPI tracing. Only specify this parameter when instructed to do so by Informatica Global Customer Support.</p>
TYPE	Yes		Specify ORCL.

Configuring the Oracle Catalog Parameters

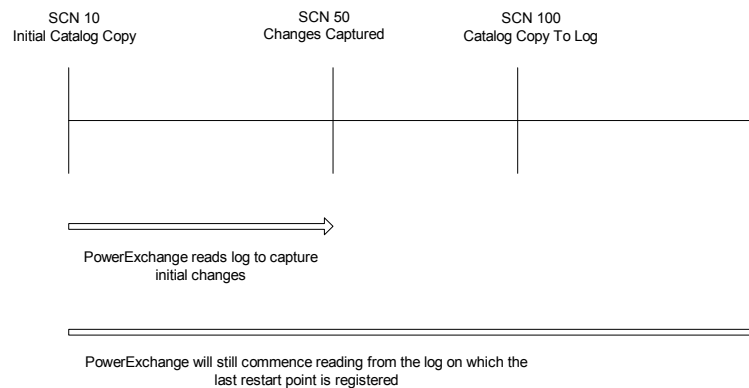
The configuration of the parameters CATINT, CATBEGIN and CATEND are important to the performance of Oracle capture. There can be an element of trial and error in the configuration because each installation can have a very different transaction profile.

For example, in the process of working through the supplied DDL in oracapt.sql, say the Oracle catalog was written to the Oracle Redo Log at SCN 10. Changes were then applied to Oracle at SCN 50. These changes

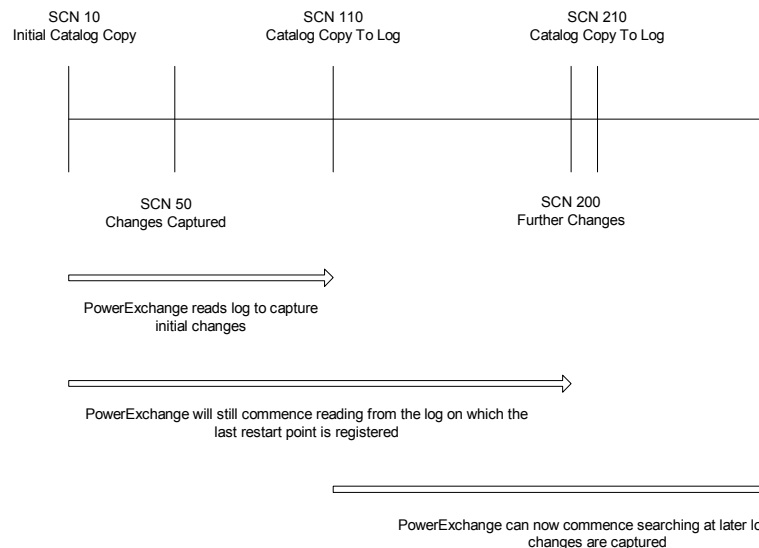
were captured by PowerExchange and the log ends at SCN 100. Additional changes are added at SCN 200. But because there is only one catalog on the log, PowerExchange needs to read from the initial catalog at SCN 10 through (but ignoring) the changes and associated restart point at SCN 50 and onwards to those changes required for capture at SCN 200. This can result in longer capture times until the next catalog is written to the Redo logs.



Now assume that you copy the Oracle log to the Redo Logs at SCN 100. This does not help because the read must start at the start of the log for the last restart point:



Now assume that the catalogs have been written to the redo logs at SCN 10, 210 and 310. When changes have been captured at SCN 50 and SCN 200, PowerExchange only needs to read from the log that includes the last restart point.



Think carefully when configuring these parameters. A copy at the default frequency of once a day might not be enough. If data is being captured continuously for a process, the catalogs need to be written often enough to prevent long restart times. If the data is captured occasionally then the catalog only needs to be captured once for each capture process.

CAPI_CONNECTION Statement for Continuous Extraction Mode

You can also configure CAPX CAPI_CONNECTION statements for continuous extraction mode from condense files.

Add the CAPX CAPI_CONNECTION statement if you use continuous extraction mode for condense files. For example:

```
CAPI_CONNECTION= (NAME=capx_name, TYPE= (CAPX, DFLTINST=collection_id) )
```

Collection_id must match the collection identifier specified in the ORACLEID statement which matches the collection identifier specified in the Registration Group containing the capture registrations.

If using continuous extraction mode for condense files, configure the CAPI_SRC_DFLT statement pointing to the UOW Cleanser CAPI_CONNECTION statement for Oracle Real-time. The Condense task uses these connections to capture changes. The Condense task does not allow CAPI_CONNECTION overrides so specification of a proper default is required. See “CAPI_SRC_DFLT Statement” on page 14.

Note: PowerExchange does not require CAPI_CONNECTION statements on the PowerExchange Navigator machine. You can register Oracle tables for CDC and perform database row test functions in Navigator using the PowerExchange Listener where Oracle CDC is running.

CAPX CAPI_CONNECTION Parameters

The CAPX CAPI_CONNECTION statement specifies the parameters used for continuous extraction from condense files. Using continuous extraction mode from condense files, extractions run near real-time using condense files as the change stream.

You can use continuous extraction mode from condense files with DB2 for Linux, UNIX, and Windows CDC, Oracle CDC, and SQL Server CDC.

The CAPX statement has the following syntax:

```
CAPI_CONNECTION= (
  NAME=name,
  [DLLTRACE=trace_id] ,
  [TRACE=trace] ,
  TYPE= (CAPX,
    DFLTINST=collection_id,
    [FILEWAIT=nnnnn] ,
    [RSTRADV=sssss]
  )
)
```

Parameter	Required	Default	Description
DFLTINST	Yes	none	Specifies the PowerExchange instance to process. For DB2 for Linux, UNIX, and Windows, specify the database from the registration group. For Oracle, specify the collection identifier from the ORACLEID statement or the instance from the registration group. For Microsoft SQL Server, specify the instance from the registration group.
DLLTRACE	No	none	Specifies the name of the TRACE= statement that activates internal DLL tracing for this specific CAPI. Only specify this parameter when instructed to do so by Informatica Global Customer Support.

Parameter	Required	Default	Description
FILEWAIT	No	1	Specifies the number of seconds to wait before checking for new condense files. Minimum is 1. Maximum is 86400.
NAME	Yes	none	Specifies a unique name for the CAPI_CONNECTION statement. Minimum is 1 character. Maximum is 8 characters.
RSTRADV	No	none	Specifies the number of seconds to wait after receiving a Unit of Work (UOW) for a source in the extraction before advancing the restart tokens by returning an "empty" UOW. Use this parameter to reduce the amount of data which is reprocessed when restarting the extraction. Default is to advance the restart tokens for a source in the extraction when a UOW for that source is read. Minimum is 0. Maximum is 86400. RSTRADV=0 returns empty UOWs regardless of whether the overall restart point has changed. It advances the restart point for every UOW encountered, regardless of whether the overall restart point for the extraction is advanced. Warning: Specifying RSTRADV=0 adversely affects performance.
TRACE	No	none	Specifies the name of the TRACE= statement that activates the common CAPI tracing. Only specify this parameter when instructed to do so by Informatica Global Customer Support.
TYPE	Yes	none	Specify CAPX.

For more information about Consumer API (CAPI) configuration parameters, see the *PowerExchange Reference Manual*.

Migrating to Condense File Extraction Modes

You can migrate real-time extractions to batch or continuous extraction modes. You can also migrate batch extractions to continuous extraction mode. Use the following rules and guidelines when you migrate real-time extractions:

- ♦ If you re-create capture registrations with PowerExchange 8.5 or later, PowerExchange Condense cannot access data that was previously condensed by a PowerExchange release earlier than 8.5. In PowerExchange 8.5, the format in which PowerExchange Condense writes change data to condense files changed to support multi-byte character sets.

The Condense task uses the capture registration to parse the changes that are written to the condense files. During capture registration creation, PowerExchange records version information in the capture registration.

Before re-creating a capture registration with PowerExchange 8.5 or later, you must first extract all captured data from the existing condense files. To prevent future extraction incompatibilities, also re-create the capture registrations that are enabled for Condense processing. After installing PowerExchange 8.5 or later, cold start the Condense task.

- ♦ Restart tokens differ between real-time, batch, and continuous extraction mode.

In order to maintain positioning within the change stream, PowerExchange constructs restart tokens. These restart tokens are used to re-establish connection to the change stream to the point of interruption when extractions are restarted.

PowerExchange creates a one format of restart tokens for real-time and batch extraction mode and a different format for continuous extraction mode. These different formats of restart tokens are not compatible. Switching an extraction from one mode to another requires a migration process.

To migrate from real-time to batch extraction mode:

1. Configure the Condense task. See “Configuring the PowerExchange Condense Task” on page 44.
2. Quiesce write activity to all tables switching from real-time to batch extraction mode.
3. Extract all captured changes and then stop the extraction.
4. Run DTLUAPPL with RSTTKN GENERATE to get restart point.

Configure DTLUAPPL to use the Oracle UOWC CAPI_CONNECTION statement. If the PowerExchange configuration file is configured with this CAPI statement as the default for Oracle, no overrides are necessary when using DTLUAPPL. Otherwise, use the CONN_OVR parameter and point to the Oracle UOWC CAPI_CONNECTION statement. See “Creating Restart Points with DTLUAPPL” on page 35.

5. Reinstate write access to all quiesced tables.
6. Delete and recreate the capture registration specifying Condense Part.
7. Update the RESTART_TOKEN and SEQUENCE_TOKEN statements in the Condense parameter file, dtlca.cfg, with restart information generated by DTLUAPPL. See “Configuring PowerExchange Condense Parameters” on page 46.
8. Cold start the Condense task. See “Starting and Stopping Condense” on page 53.
9. Change the extraction from real-time to batch extraction mode.

If using PWXPC, change the session to use a PWX Oracle CDC Change connection instead of a PWX Oracle CDC Real Time connection. See the *PowerExchange Interfaces for PowerCenter*

If using ODBC, update the ODBC data source changing the DBTYPE from CAPXRT to CAPX. See the *PowerExchange Reference Manual*.

10. Cold start the PowerCenter extraction session.

To migrate from real-time to continuous extraction mode:

1. Configure the Condense task. See “Configuring the PowerExchange Condense Task” on page 44.
2. Quiesce write activity to all tables switching from real-time to continuous extraction mode.
3. Extract all captured changes and then stop the extraction.
4. Run DTLUAPPL with RSTTKN GENERATE to get restart points.

Configure DTLUAPPL to use the Oracle UOWC CAPI_CONNECTION statement. If the PowerExchange configuration file is configured with this CAPI statement as the default for Oracle, no overrides are necessary when using DTLUAPPL. Otherwise, use the CONN_OVR parameter and point to the Oracle UOWC CAPI_CONNECTION statement. See “Creating Restart Points with DTLUAPPL” on page 35.

5. Reinstate write access to all tables that were quiesced.
6. Delete and recreate the capture registration specifying Condense Part.
7. Update the RESTART_TOKEN and SEQUENCE_TOKEN statements in the Condense parameter file, dtlca.cfg, with restart information generated by DTLUAPPL for the Condense task. See “Configuring PowerExchange Condense Parameters” on page 46.
8. Update the PowerExchange configuration file, dbmover.cfg, with the CAPX CAPI_CONNECTION. See “CAPI_CONNECTION Statement for Continuous Extraction Mode” on page 19.
9. Cold start the Condense task. See “Starting and Stopping Condense” on page 53.
10. Change the extraction from real-time to continuous extraction mode.

If using PWXPC, update the PWX Oracle CDC Real-time connection in the session to specify the name of the CAPX CAPI_CONNECTION statement in the CAPI Connection Name Override attribute. See the *PowerExchange Interfaces for PowerCenter*.

If using ODBC, override the ODBC data source with the name of the CAPX CAPI_CONNECTION statement using the SQL Escape Sequence DTLCONNOVR. See the *PowerExchange Reference Manual*.

11. Cold start the PowerCenter extraction session.

To migrate from batch to continuous extraction mode:

1. Stop the Condense task. See “Starting and Stopping Condense” on page 53.
2. Extract all captured changes and then stop all extractions.
3. Run DTLUAPPL with RSTTKN GENERATE to get restart point.

Configure DTLUAPPL to use the Oracle UOWC CAPI_CONNECTION statement. If the PowerExchange configuration file is configured with this CAPI statement as the default for Oracle, no overrides are necessary when using DTLUAPPL. Otherwise, use the CONN_OVR parameter and point to the Oracle UOWC CAPI_CONNECTION statement. See “Creating Restart Points with DTLUAPPL” on page 35.

4. Update the RESTART_TOKEN and SEQUENCE_TOKEN statements in the Condense parameter file, dtlca.cfg, with restart information generated by DTLUAPPL for the Condense task. For more information, see “Configuring PowerExchange Condense Parameters” on page 46.
5. Delete and recreate the capture registration specifying Condense Part.
6. Update the PowerExchange configuration file, dbmover.cfg, with the CAPX CAPI_CONNECTION. See “CAPI_CONNECTION Statement for Continuous Extraction Mode” on page 19.
7. Cold start the Condense task. See “Starting and Stopping Condense” on page 53.

Note: The Condense task deletes all existing condense files when cold started. You must ensure that all condensed data for all table is extracted before doing the cold start.

8. Change the extraction from real-time to continuous extraction mode.

If using PWXPC, change the session to use a PWX Oracle CDC Real Time connection instead of a PWX Oracle CDC Change connection. Specify the name of the CAPX CAPI_CONNECTION statement in the CAPI Connection Name Override attribute. See the *PowerExchange Interfaces for PowerCenter*.

If using ODBC, update the ODBC data source changing the DBTYPE from CAPX to CAPXRT. Override the ODBC data source with the name of the CAPX CAPI_CONNECTION statement using the SQL Escape Sequence DTLCONNOVR. See the *PowerExchange Reference Manual*.

9. Cold start the PowerCenter extraction session.

Oracle Datatypes

PowerExchange uses Oracle LogMiner to retrieve changes from the Oracle redo logs. Oracle does not log or fully log certain datatypes in the Oracle redo logs.

The following table lists Oracle datatypes and indicates whether PowerExchange supports the datatype for change data capture:

Table 3-5. Oracle Data Types and Oracle CDC

Datatype	Supported?	Notes
BFILE	No	Not fully logged in the Oracle redo logs.
BINARY_DOUBLE	Yes	Requires Oracle client 10.2.0.1 or higher
BINARY_FLOAT	Yes	Requires Oracle client 10.2.0.1 or higher
CHAR	Yes	

Table 3-5. Oracle Data Types and Oracle CDC

Datatype	Supported?	Notes
DATE	Yes	
FLOAT	Yes	
LOB	No	
LONG	No	
LONG RAW	No	
NUMBER	Yes	PowerExchange handles NUMBER columns as follows: - Numbers with scale of 0 and precision less than 10 are treated as INTEGER. - Numbers with a defined precision and scale are treated as NUMCHAR. - Numbers with an undefined precision and scale are treated as DOUBLE.
RAW	Yes	
TIMESTAMP	Yes	
TIMESTAMP WITH TIME ZONE	No	Not fully logged in the Oracle redo logs.
TIMESTAMP WITH LOCAL TIME ZONE	No	Not fully logged in the Oracle redo logs.
VARCHAR2	Yes	

SQL Loader Restrictions

Oracle CDC supports capture of data loaded by SQL Loader with the following restrictions:

- ♦ The load must be conventional path. Direct path loads are not captured by LogMiner, and therefore cannot be captured by Oracle CDC.
- ♦ The load method should be either Insert, Append, or Replace. Truncate causes the TRUNCATE TABLE DDL statement to be issued. Oracle CDC does not capture DDL, so it would not capture the row deletions done by the TRUNCATE TABLE statement.

Performance Considerations

Consider the following when configuring and using Oracle CDC:

- ♦ Use real-time extraction mode only when running very few concurrent extractions. PowerExchange Oracle CDC creates a LogMiner session for each real-time extraction. LogMiner sessions are resource intensive and impact overall system responsiveness. Use continuous extraction mode from Condense files instead.
- ♦ If using continuous extraction mode, minimize the size of the CDCT. Oracle CDC reads the CDCT each time the FILEWAIT interval expires. Large CDCT files increase the I/O required to read the CDCT thereby impacting system resources and increasing extraction latency.

Manage the CDCT size by specifying a reasonable value in the Condense configuration statement COND_CDCT_RET_P. The Condense task uses the value in this statement to delete old condense files and their related CDCT entries.

CHAPTER 4

Registering CDC Sources

This chapter includes the following topics:

- ◆ Introduction, 25
- ◆ Registering Sources for Change Data Capture, 25
- ◆ Oracle Registration Process, 26
- ◆ Extraction Maps, 29
- ◆ Data Delivery Format, 34
- ◆ Creating Restart Points with DTLUAPPL, 35
- ◆ Application Group, 38

Introduction

This section describes the process for registering an Oracle data source with the PowerExchange Navigator. Oracle data can be captured from Linux, UNIX, and Windows environments. For more information about the PowerExchange Navigator, see the *PowerExchange Navigator User Guide*.

Registering Sources for Change Data Capture

PowerExchange Oracle CDC collects and extracts changes only when the source is properly registered. In the PowerExchange Navigator, register the tables for which you wish to use PowerExchange Oracle CDC. The registration process consists of using the PowerExchange Navigator as well as executing Oracle SQL. Prior to registering tables for Oracle CDC, configure Oracle and PowerExchange. See “Configuring Oracle Change Data Capture” on page 7.

After you configure PowerExchange and Oracle, perform the following tasks to create a capture registration:

1. Add a Registration Group in the PowerExchange Navigator, if one does not already exist. See “Oracle Registration Process” on page 26.
2. Register the source using the PowerExchange Navigator.
3. Establish a capture start point using DTLUAPPL. See “Creating Restart Points with DTLUAPPL” on page 35.
4. Set the PowerExchange capture registration to Active.

Option	Description
Type	Select Oracle.
User ID and Password	User ID and password to be used when accessing the location and instance. The Oracle user ID must have access to the Oracle location mentioned above through the Navigator. If you do not enter a required password, the Navigator prompts for the password.
Collection Identifier	This must match the Collection_id specified in the ORACLEID statement in the dbmover.cfg. See "ORACLEID Statement" on page 13.
Add Registration	Identifies whether the capture registration details will be provided now or later. If the option is not selected, the process completes at this point and the Next button changes to Finish.

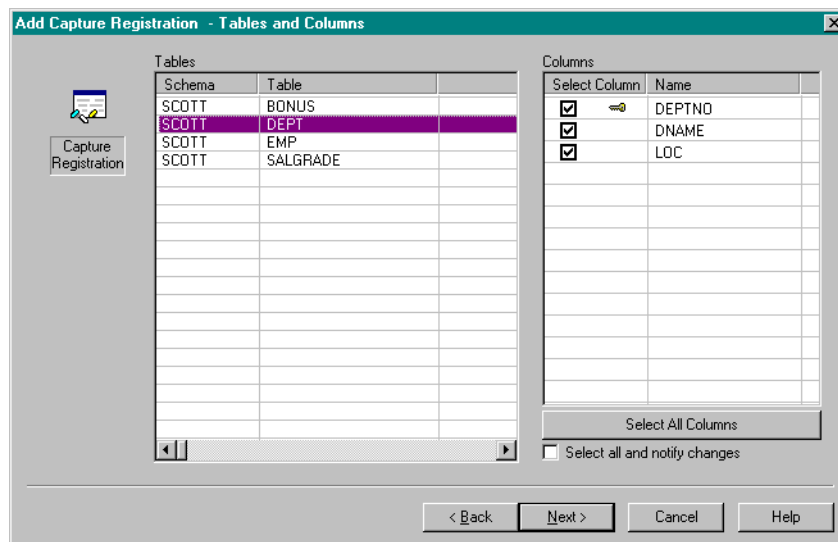
3. Click Next.
4. To filter the tables returned when selecting a source table, enter the following information in the Add Capture Registration - Name and Table Filter dialog box:

Option	Description
Name	Name of this registration within the Registration Group. This is limited to eight alphanumeric characters.
Table Filter	Minimizes the actual selection list retrieved from the specified location and instance.
Schema	Source data creator or schema name. Filter where possible, otherwise a large amount of information may be returned.
Table	The required table, if known.

5. Click Next.

A list of source tables appears. This dialog box defines the actual source being registered for data capture. A capture registration can only contain one table.

There can only be one capture registration for a given source. Take care when selecting the individual columns for capture. An extraction process can select a subset of the captured columns, but only for columns that have been captured.



The display window shows a list of tables. When you select a table, its columns display on the right. You can select or clear a column for capture.

Columns marked with a key are the index columns for that table. If the table does not have any key columns, you must capture all columns in the table.

Note: Changes to the table structure after registering the columns require updating the capture registration. You can use versioning to handle capture registration updates.

6. Select the table and some or all of the columns, and click **Next**.
7. Enter the following information to determine the type and characteristics of the capture registration.

Option	Description
Type	Synchronous is currently the only capture registration type for Oracle capture.
Condense	Condense type: <ul style="list-style-type: none">- None. Indicates that condense files are not kept. Use for CAPXRT capture type only. It does not require condense processing and does not create condensed files.- Part. Keeps all committed updates for a given row in the source.
Status	Capture registration status: <ul style="list-style-type: none">- Inactive. The registration is to be defined but is not to be used by the Collection Agent.- Active. The registration is effective immediately.
Supplement Log Group Name	Name of the supplemental log. Completing the registration results in the automatic generation and execution of a DDL to enable supplemental logging. Without Supplemental Logging enabled, Oracle capture will not work.
Comment	Description of the capture registration.
Execute DDL now	Registration generates a DDL to enable supplemental logging. You can run the DDL immediately or store and run it later. You are prompted to name the DDL file. The user created within the ORACAPT.SQL file will not have the authority required to run this statement.

8. Click **Finish**.

Note: The Registration is written to Capture Registration Repository file on the source platform.

Adding a Capture Registration to an Existing Group

To add a capture registration to an existing group:

1. Right-click the required Registration Group and click **Open**.
2. Click the **Add Capture Registration** icon on the toolbar.

The Add Capture Registration - Name and Table Filter dialog box appears. Follow steps 3 to 8 on page 27.

Deleting a Registration Group

Deleting a Registration Group does not delete registration entries. It only deletes group information held locally on the workstation.

To delete a registration group:

1. Select the appropriate Registration Group and click the **Delete** icon.
A confirmation dialog box appears.
2. Click **Yes**.

If there are no active registrations outstanding for this group, the delete is accepted and the group is removed. Otherwise a warning message appears. The warning is intended to prevent a group from being deleted while an active registration continues to capture data.

Deleting a Registration Entry

To delete a registration entry:

- ▶ Select the appropriate Registration Entry and click the **Delete** button.

If the registration is not active, you are asked to confirm the deletion, otherwise you are prompted to make the registration history before deleting. This deletes the registration entry on the source platform. Data will no longer be picked up by the extraction process.

Viewing a Registration Group

A complete list of registrations is displayed under the Registration Group in the Resource Explorer.

To view a registration group:

- ▶ Select the registration group in the Resource Explorer.

The details for the group appears in the Resource Inspector.

Viewing a Capture Registration

To view the details of a capture registration:

- ▶ Double-click a capture definition group on the Resource Explorer Resource tab or select a capture registration from the group list. Then double-click the registration that you want to view.

The following information appears:

- ♦ **Resource Explorer.** Displays a list of capture registration groups and the capture registrations within each group. To select and open a capture registration double-click the Capture Registration Name. All the Registration details are displayed in the Capture Registration details dialog boxes.
- ♦ **Capture Registration details.** Displays the column information of the table referenced by the capture registration. This only displays details for those columns that are selected for capture. This includes:
 - Name
 - Table name
 - Column number
 - Type
 - Length
 - Scale
 - Key

To make changes to an existing list of columns for a given source, right-click the registration name.

- ♦ **Resource Inspector.** Displays the registration properties of the capture registration. For example, Type will be Synchronous and the Tag contents will be used to access the registration data in the Repository.

If you make any changes, they must be confirmed by clicking Apply. Once confirmed, the Registration in the Explorer changes color. This is an indication that the source needs to be saved.

Extraction Maps

An extraction process is a process of extracting changes from PowerExchange Change Capture. The extraction type used for retrieving captured Oracle data is CAPXRT. As part of the capture registration process, an extraction group with the same name is automatically defined. You can modify the default definitions.

Note: You should use a separate application for each extraction. If you have multiple extractions in a single application, each extraction must be run serially.

If additional extractions processes are required, you must define them by using the Add Extract Definition wizard. When extracting changes for a source, the user, PowerCenter or third-party software, specifies a standard SQL request in the following format:

```
Select * from DataMap.Name
```

This request is the same as any other SQL request given to the PowerExchange API. When using the default data map that was created as part of the capture registration process, the request that PowerExchange generates has the following format:

```
Select * from Dxcollectionid.creator_tablename
```

Where:

- ♦ D is the default entry, or U for all user-defined data maps.
- ♦ *x* is the value 8 for Oracle sources.
- ♦ *collectionid* is the collection identifier that you specified when creating the registration group.
- ♦ *creator_tablename* is the creator ID and table name of a source table.

Defining the Extraction Process

The extractions are defined under the Extraction Group under Resources in the Resource Explorer. When a Registration Group is defined, an Extraction Group with the same name is automatically created by default. There is generally no need to define Extraction Groups manually.

Adding an Extraction Group

To add an extraction group:

1. On the toolbar, click the Add Extraction Group icon, or right-click Extraction Group and click **Add Extraction Group**.
2. Enter the following information:

Item	Description
Name	The name of the extraction group.
Location	Specifies the name as defined in the dbmover.cfg file where the PowerExchange Listener is located.
Type	The type of data to be extracted.
UserID/Password	User ID and password to identify the user and check their authorizations.
Collection Identifier	Collection identifier as specified in the ORACOLL parameter of the ORCL CAPI_CONNECTION statement in dbmover.cfg.
Add Extraction Definition	Indicates if the extraction details will be given now or later. If the box is cleared, the process completes at this point and the Next button changes to Finish. When you click Finish, all existing Extraction Definitions for the specified Type, Database Instance and Location are displayed.

3. To add an extraction definition, click **Add Extraction Definition** and click **Next**.

4. Enter the following information.

Item	Description
Schema Name	A fixed name based on the following naming convention: <ul style="list-style-type: none">- U. User defined.- DB type. 8 for Oracle.- DB instance. Such as oemcap. The schema name is unique within a specific listener, database instance and database type.
Map Name	A name given to that specific extraction definition. The default entry contains the registration name. Enter up to eight characters for a user extraction definition.
Table Name	A name given to the table part of the data map name. The default entry contains the actual table name. Enter up to 20 characters to define a user extraction definition.

The constructed data map name has the format:

`Uldbid.Mapname_TableName`

5. Click Next.

The Add Extract Definitions - Capture Registrations dialog box appears. With this dialog box, you can add a version of a capture registration to the extraction process definition.

The Available Capture Registrations list displays the capture registrations available for selection. The Current Capture Registrations list displays the name of the capture registrations that have been selected for the extraction definition.

You can use this to manage the versions of registrations associated with the extraction process. For example, you could remove a version 1 registration and then add the next version.

The Available Capture Registrations area displays the following information:

Field	Description
Name	Displays the name of the capture registrations that are available for selection.
Version	Displays the version number of the capture registration.
Condense	Displays the condense state of the capture registration. Must be None at the present time.
Status	Displays the status of the capture registration. This can have one of two values: <ul style="list-style-type: none">- I. Inactive. The registration is to be defined but is not to be used by the Collection Agent.- A. Active. The registration is to be defined and used by the Collection Agent.

The Current Capture Registration area displays the following information:

Field	Description
Name	Displays the name of the capture registration that has been selected.
Version	Displays the version number of the capture registration.
Condense	Displays the condense state of the capture registration.
Status	Displays the status of the capture registration.

6. Use the Add and Remove buttons to define the capture registrations used by the extraction definition.

Specify Extraction Criteria

There are a variety of values which can be specified when running an extraction process. When using the ODBC interface to PowerExchange, these criteria have to be defined as part of the ODBC definition. ODBC allows various overrides using the following escape sequences:

Parameter	Value	Parameters	Description
DTLAPP		Name	Allows you to specify the capture extract application ID without having to define a new data source when accessing PowerExchange (CAPXRT) data through ODBC.
DTLDSN		DSN	Allows SQL to override the physical file name coded in the data map. This enables you to use ODBC to read NRDB and NRDB2 data from various files through one ODBC data Source. For ODBC, this is coded anywhere in the SQL as {DTLDSN=xxx}.
DTLXTYPE/Type	SL	None	Extract all new captured data since the last extraction.
	RS	None	Default DTLXTYPE. Run the last-specified extraction process again. This is either the last extraction that was run or the one specified with the Reset Start Point Wizard in the Application dialog box. For more information about Application Groups, see "Application Group" on page 38.

For more information, see the *PowerExchange Reference Manual*.

Adding an Extraction Definition to an Existing Group

To add an extraction definition to an existing group:

1. Right-click the extraction group you want to use and click **Open**.
2. On the toolbar, click the **Add Extract Definition** icon.
The Add Extract Definition - Name dialog box appears.
3. To add the extraction definition, continue with step 4.

Deleting an Extraction Group

Deleting an extraction group deletes group-related information. It does not delete extraction definitions.

To delete an extraction group:

1. Select the Extraction Group you want to delete and click the **Delete** icon.
A confirmation dialog box appears.
2. Click **Yes**.

Deleting an Extraction Definition

After you delete an extraction definition, you cannot extract any changes using that extraction definition. Be sure to make corresponding changes to avoid invalidating data.

To delete an extraction definition:

- Select the extraction definition you want to delete and press the **Delete** button.

Viewing an Extraction Definition

To view extraction definition details:

1. Double-click an extraction group from the Resource Explorer Resource tab.
2. Select an extraction definition from the group contents and double-click on the Registration Name.

The following information appears:

- ♦ **Resource Explorer.** Displays a list of Extraction Groups and the Extraction Definitions within each group. To select and open an extraction definition, double-click on the Extraction Name. All the Extraction details are displayed in the Extract Definition details panels.
- ♦ **Extract Definition.** It is possible to remove columns from the extraction definition by clearing the associated box. By default, automatically generated columns are not displayed. You can right-click the Table Details area and select Show Auto Generated Columns from the pop-up menu. The following columns appear:

Column	Description	Datatype	Length
DTL__CAPXRESTART1	Coded position of the extracted data in the extraction log.	VARBIN	255
DTL__CAPXRESTART2	Coded start point for the oldest unit-of-work.	VARBIN	255
DTL__CAPXUOW	Coded position of the start of the commitment control for the unit of work for this update.	VARBIN	255
DTL__CAPXUSER	If available from Oracle LogMiner, the user that performed the change. The value may also be null.	VARCHAR	255
DTL__CAPXTIMESTAMP	The time of the change to the Oracle data in the format: YYYYMMDDHHMMSS. Note that Oracle does not support microseconds in the timestamp.	CHAR	20
DTL__CAPXACTION	The change capture action: - I. Insert. - D. Delete. - U. Update.	CHAR	1
DTL__BI	The Before Image of the changed field is also retrieved in addition to the normal retrieval of the After Image. For more information about DTL__BI, see the <i>PowerExchange Navigator User Guide</i> .	Datatype and length of the source column.	
DTL__CI	A change indicator, CI, can be added to each column. For updates this indicates on retrieval whether the column has been changed, Y for yes, N for no. Inserts and Deletes return a null value. For more information about DTL__CI, see the <i>PowerExchange Navigator User Guide</i> .	CHAR	1

You can select these columns as part of the extraction definition. Click the File > Save to save any changes.

- ♦ **Extraction details.** Displays the name, version and referenced table of the extraction definition.
- ♦ **Table details.** Displays the column information of the table referenced by the Extraction Definition. This includes:
 - Name
 - Type
 - Precision
 - Scale
 - Length

In addition, null indicator hidden columns are shown adjacent to nullable columns and a length indicator when these are of variable length.
- ♦ **Resource Inspector.** The Extraction Group tab displays the connection properties of the Extraction definition. If you make any changes, click **Apply** to confirm them.

Data Delivery Format

Oracle captures the Before Image, the image before the latest change, the After Image, and the latest change for updating the Oracle tables.

The way this captured data is delivered to the end user application depends upon the Image Type defined in:

- ♦ PWX ORACLE CDC application connection Image Type attribute in PowerCenter
- ♦ CAPXImageType through ODBC
- ♦ SQL Escape Sequence DTLIMTYPE

The following table describes the different ways captured update data can be delivered:

Image Type	Description
BA*	Delivers the Before Image and the After Image. The Before Image is shipped with an action character forcing the deletion of the record. The After Image, which includes changed data, is in the format of an Insert.
AI	Delivers only the After Image data for the update.
TU*	Delivers the data Before and After Images as they occur. The Before Image enables the application to ensure nothing has changed in that record since the request. The After Image appears as a change, that is, as an Update record. The update is a full image, but the application picking this up can change the target record rather than Delete it and Insert the new one as when using BA. This option is not available with PWX ORACLE CDC application connections.
<i>Note: * Inserts and deletes are delivered without Before Images.</i>	

The choice of data extraction types is an important decision and is dependant both upon the manner in which updates are applied to the target database, and on the amount of data that is to be processed. An Image Type of AI delivers After Images only.

The database permits updates of primary keys so it must be remembered that if this is allowed by applications to the source data then, if the Image Type is AI, only an After Image will be delivered as the updated image and there will be no before image to enable the application to position the update.

Where primary keys may be updated in the source data, then selecting BA is the safe option. This will deliver the before image to be processed as a delete and the after image will then be processed as an insert.

Creating Restart Points with DTLUAPPL

The DTLUAPPL utility creates a baseline point to which data capture can be aligned. Each subsequent extraction registers new sequence and restart tokens. There are two ways to create a restart point for a capture extraction:

- ♦ Using the utility DTLUAPPL. See the *PowerExchange Utilities Guide*.
- ♦ Running an extraction without specifying a restart point so the default restart point is used. For information about the default restart points, see *PowerExchange Interfaces for PowerCenter*.

If using PWXPC connections in PowerCenter, for more information about managing restart tokens, see *PowerExchange Interfaces for PowerCenter*.

If using ODBC, the restart tokens can be viewed and manipulated through the PowerExchange Navigator. If a data capture process needs to be run again, you can reset the restart tokens to an earlier point in time. When you do this, make sure all relevant captured changes are still available.

The DTLUAPPL utility is designed to set or reset start points for the extraction process. DTLUAPPL uses the following files:

- ♦ **dtltnp.txt**. Contains with sample configurations and is the default file name for DTLUAPPL statements.
- ♦ **Log file**. Contains information appended by DTLUAPPL.

DTLUAPPL can both register and modify applications, and then print the status of those applications with its print utility.

Running DTLUAPPL in Windows

To run DTLUAPPL in Windows, edit `dtltnp.txt` to create or modify the applications as required and run DTLUAPPL from a command prompt. PowerExchange provides sample statements in `dtltnp.txt` in the PowerExchange install directory.

There are three ways of running DTLUAPPL:

- ♦ Issue the DTLUAPPL command on the command line. The utility assumes the tokens are in `dtltnp.txt` and displays the output in the command window.
- ♦ Issue the following command:

```
DTLUAPPL > logname.txt
```

The utility reads the statements from `dtltnp.txt` and writes the output to `logname.txt`.

- ♦ Issue the following command:

```
DTLUAPPL myfile.txt > logname.txt
```

The utility reads the statements from `myfile.txt` and writes the output to `logname.txt`.

Running DTLUAPPL on Linux and UNIX

Edit `dtltnp.txt` to create or modify the applications as required and run DTLUAPPL from a command prompt. PowerExchange provides sample statements in `dtltnp.txt` in the PowerExchange install directory. There are three ways of running DTLUAPPL:

- ♦ Issue the DTLUAPPL command on the command line. The utility reads the statements from `dtltnp.txt` and displays the output in the command window.
- ♦ Issue the following command:

```
dtluappl > logname.txt
```

The utility reads the statements from `dtltnp.txt` and writes the output to `logname.txt`.

- ♦ Issue the following command:

```
dtluappl myfile.txt > logname.txt
```


00000000000000000000500064F454D524550
END APPL ORAAP1

Example 2 - Generate Restart Tokens for Continuous Extraction Mode

In this example, an application called `dummy` is modified with restart tokens for registration `rrtb001`. The default `CAPX_CONNECTION` in `dbmover.cfg` is overridden with a `CAPX_CONNECTION`. The capture method, `CAPTMETHOD`, is `CAPXRT` and the condense type, `CONDTYPE`, is `P` for partial.

DTLUPPL generates restart tokens for continuous extraction mode when the CAPI_CONNECTION statement used is TYPE=CAPX and the CAPTMETH=CAPXRT.

```

UID user01 EPWD 40ABC4B0E32FD99F CONN_OVR CAPXORA
MOD APPL dummy FOX920 RSTTKN GENERATE CAPTMETH=CAPXRT CONDTYPE=P
MOD RSTTKN rrtb001
END APPL dummy
PRINT APPL dummy

```

Example 3 - Generate Restart Tokens for Remote Database

The following example shows how an application can be added on a remote database:

```
ADD APPL ORAAP3 ORACLEID RSTTKN GENERATE ORACONN=OCONN ORAINST=OINST ORACOLL=OCOLL
ADD RSTTKN oraemp2
END APPL ORAAP3
```

Modifying an Existing Application

The syntax for the modify application is similar to adding an application, using “MOD” instead of “ADD.” If the application to be modified does not exist, a warning is issued and the application is added. An application may be modified to allow the application/data load to be run again from an earlier point in time.

The SEQUENCE and RESTART tokens can be set explicitly in a modification by using the following syntax:

[illegible]

Note: In dtltknp.txt each of these tokens fits in a single line.

For descriptions of the DTLUAPPL parameters, see “Adding a New Application” on page 36.

Printing an Application

The DTLUAPPL utility can print applications, providing the status of existing sequence and restart information. When running on a Windows environment, run this utility from a command line where the output will be displayed. The syntax of the print command is:

```
PRINT APPL appname | ALL
```

This generates details for either the application or all applications, as directed.

You can print more than one application by concatenating several PRINT APPL commands. You can combine ADD and MOD actions with PRINT commands in a single request, as follows:

[illegible]

Application Group

Application groups provide the ability for you to display the information for a given extract application and to allow you to reset the status of the extractions process to run again.

The Application Group is based on the same principle as the Extraction and Registration Groups. The Group level identifies the location of the PowerExchange Listener task. There is no need for a database instance type or name. The Resource Inspector displays the following:

- ◆ Name
- ◆ Location
- ◆ User ID
- ◆ Password

The Application Name list is based on the application name on the extraction run.

Click on an Application Name to display a list of all the extracts in timestamp order, latest one on top. Click on the timestamp to display the complete information for that extraction run. This includes information such as the registration name, table name, and CND files read.

Reset to New Start Point

Use Reset to New Start Point to reset the start point of the next extraction run for the selected application. It resets the start point to the start point of the open timestamp.

The application can then either:

- ◆ Restart at the start of the selected run
- ◆ Be reset again

To restart a new start point:

1. Open a timestamp and review the associated extract.
2. Right-click on the application and select Reset to New Start Point.

Note: This cannot be undone. If in doubt, use DTLUAPPL. Tokens can be printed for copy and paste using DTLUCDEP.

CHAPTER 5

Amending CDC Definitions

This chapter includes the following topics:

- ♦ Overview, 39
- ♦ PowerExchange View of the Captured Data, 39
- ♦ Changing the Underlying Data Source, 40
- ♦ Changing Capture Definition Properties, 40

Overview

After you register a data source for change capture, you might need to edit the registration account for changes to your processing or underlying data.

Any change has implications for the PowerExchange change capture process. The following major changes are allowed:

- ♦ Updating PowerExchange view of the captured data
- ♦ Changing the underlying data source
- ♦ Changing properties of the registration definition

PowerExchange View of the Captured Data

The simplest way to update the PowerExchange view of underlying data is to remove columns from the extraction map. By leaving the registration unaltered, all the data from the data source is still captured, but PowerExchange only extracts the data in the columns remaining in the extraction map.

Because source data for removed columns are still captured, you can add the columns to the extraction map at any time to enable PowerExchange to extract data from those columns again.

You do not need to restart the change capture process or the source system after performing this change.

Removing Columns from the Extraction Map

You can remove a column from an extraction map by clearing the check box for any columns you no longer require.

This does not prevent the data being captured. You can select the column again to make data available for extraction through this extraction map.

Changing the Underlying Data Source

If changes are made to an Oracle table registered for capture, use the following steps to allow the new format of data to be captured while retaining the ability to access historically captured data.

Managing Changes to the Data Structure

This section describes how to manage schema changes for the source tables that are involved in change propagation.

To change an Oracle source schema:

1. Ensure that all changes that occurred under the original schema have been extracted and applied to the target.
2. Change the Oracle table.
3. In PowerExchange, create a new registration to reflect the changes to the table. For more information, see “Registering Source Data” on page 26.
4. If using Condense, recycle the Condense task to pick up the new registration. See “Starting and Stopping Condense” on page 53.

Changing Capture Definition Properties

You can change some of the properties of the capture registration or extraction map. You can amend the following properties of the capture registration:

- ♦ Status
- ♦ Condense
- ♦ Comments

For example, you might have created a registration with a status of Inactive and now wish to use it in your environment. You can change the status to Active. If a registration is no longer required, you can disable the registration by changing the status to History.

When you amend the capture registration, the version indicator is increased. Change the extraction map to use the new capture registration version.

To amend a capture registration:

1. Select the relevant registration group and open the capture registration.
2. Click the capture registration you want to amend.

Changes cannot be made to the columns selected in this screen, although columns can be removed from the extraction definition.

3. Change the status, Condense, or add comments for the capture registration.

Changing Extraction Definition Properties

To amend an extraction definition:

1. On the Extraction Group tab of the Resource Explorer, right-click the extraction group containing the extraction definition you want to amend and click Amend Capture Registrations.
2. Select the capture registration you want to amend.
3. The Available Capture Registrations area contains the following information:

Item	Description
Name	Displays the name of the capture registrations that are available for selection.
Version	Displays the version number of the capture registration.
Condense	Displays the condense state of the capture registration.
Status	Displays the status of the capture registration. This can have one of two values: <ul style="list-style-type: none">- I. Inactive. The registration is to be defined but is not to be used by the Collection Agent.- A. Active. The registration is to be defined and used by the Collection Agent.

The Current Capture Registration area contains the following information:

Item	Description
Name	Displays the name of the capture registration that has been selected.
Version	Displays the version number of the capture registration.
Condense	Displays the condense state of the capture registration.
Status	Displays the status of the capture registration.
Add	Adds the selected capture registration to the Current Capture Registrations list.
Remove	Removes the selected capture registration from Current Capture Registrations list.

CHAPTER 6

PowerExchange Condense

This chapter includes the following topics:

- ◆ Introduction, 43
- ◆ Configuring the PowerExchange Condense Task, 44
- ◆ Configuring PowerExchange Condense Parameters, 46
- ◆ Configuring Flexible Condense, 51
- ◆ Starting and Stopping Condense, 53

Introduction

PowerExchange Condense accesses changes from flat files, called condense files, which are created by the condense subtask. The PowerExchange Condense task uses Oracle LogMiner to read the Oracle redo logs and capture the changes to the condense files. By using the Condense CAPTPARM parameters, you can control the frequency with which the data is condensed.

PowerExchange accesses the change data in condense files by using one of the following modes:

- ◆ Batch extraction mode for condense files using the CAPX access method
- ◆ Continuous extraction mode for condense files using the CAPXRT access method

Registering for Condense

PowerExchange Condense only condenses data for capture registrations that specify condense processing. You specify condense processing for a capture registration by:

- ◆ Specifying Part for the Condense option in the PowerExchange Navigator capture registration
- ◆ Activating the capture registration

See “Registering Sources for Change Data Capture” on page 25.

Condense uses only active capture registrations that have a Condense type of Part. If no registrations meet these criteria, the Condense task reports that there are no active registrations and ends.

For more information about modifying existing registrations to take advantage of the Condense option, see “Amending CDC Definitions” on page 39.

Migration Consideration

If you created capture registrations for tables by using a PowerExchange version earlier than 8.5, PowerExchange will not be able to access any data that was condensed for those tables.

The Condense task uses the capture registration to parse the changes and to write that data to the condense files. When you created the capture registration, PowerExchange recorded version information as part of the registration.

Condense writes the changes to the condense files in a different format if the capture registration was created by using a PowerExchange version earlier than 8.5. This format change allows Condense to support multi-byte character sets. Before re-creating a registration, you must first extract all captured data from existing condense files. To prevent any future extraction incompatibilities, re-create any registrations that are enabled for Condense and cold start the Condense task after installing PowerExchange 8.5 or later.

Configuring the PowerExchange Condense Task

PowerExchange Condense is comprised of the following subtasks:

- ♦ **Controller.** This is the job step task and controls the address space and starts the subtasks.
- ♦ **Condense.** This subtask is specifically responsible for condensing data.
- ♦ **Command Handler.** This subtask provides the command interface to the Condense task.
- ♦ **Dump.** This subtask provides dump services to the Condense task.

The PowerExchange log contains messages indicating when the various tasks start and end and, generally, from which task a message is being issued.

PowerExchange Condense Operational Modes

The Condense task has two modes of operation:

- ♦ Batch
- ♦ Continuous

Set the mode of operation using the `COLL_END_LOG` parameter. For more information, see “Configuring PowerExchange Condense Parameters” on page 46.

PowerExchange Condense Batch Mode

Batch mode means that a single condense cycle is run and then the Condense task shuts down. Running the Condense task in this fashion is well suited to batch applications. For example, single condense runs might be inserted at appropriate points in an automated schedule after update jobs.

Specify batch mode for the Condense task by coding `COLL_END_LOG=1` in the Condense configuration file.

PowerExchange Condense Continuous Mode

Continuous mode means that the condense task runs continuously for an extended period of time. In this mode, the Condense task “sleeps” after each condense cycle. The next condense cycle is triggered by one of the following:

- ♦ The number of minutes defined by parameter `NO_DATA_WAIT` is reached.
- ♦ A `CONDENSE` command is manually entered.
- ♦ A `FILESWITCH` command is manually entered.
- ♦ A `SHUTCOND` command is manually entered.

In Continuous mode, the Condense task does not shut down. You must shut it down using the SHUTDOWN command from the command line.

File switch processing closes the open files if they contains data and opens a new set of condense files for future changes. The Condense task automatically does a file switch when the criteria defined by the FILE_SWITCH_CRIT, FILE_SWITCH_VAL, and the FILE_SWITCH_MIN configuration parameters are met. File switch also occurs if you manually issue a FILESWITCH command.

Note: The Condense task does not file switch if no data is present in the current condense file. The file switch is skipped until the next interval. The file switch continues to be skipped until an interval expires and there is data in the condense file.

Specify continuous mode for the Condense task by coding COLL_END_LOG=0 in the Condense configuration file.

PowerExchange Condense Output Files

The Condense task writes output files to the directory specified by the CAPT_PATH parameter in the PowerExchange dbmover.cfg configuration file. If no CAPT_PATH parameter is specified, the Condense task writes the output files to the directory from where it is invoked.

CDCT File

The CDCT file is updated by the Condense task. The PowerExchange Listener reads the CDCT when extracting data.

After each file switch, the Condense task writes records to the CDCT containing information about each condense file. These keyed records contain information about the data that has been condensed such as the condense file name, the start and end times, whether before images are included, the number of records in the file, and other control information.

For example, if a partial condensed file contains records for two registration tags, then the Condense task writes three records to the CDCT. All records show the name of the file. Two of the records show the record count for each registration tag. The other record has no registration tag and counts the number of “commit” records in the file, one for each unit of work.

CDCT records match DCT records on the Checkpoint file. Each time the Condense task s warm started, it checks to ensure the records match. If they do not match, the Condense task adjusts the CDCT by either inserting or deleting records as appropriate.

Condense Files

The Condense task creates condense files during the condense cycle if there are captured changes to write. The condense files contain the change data for the active registrations the Condense task found during its initialization. The Condense task customizes the names of the condense files using the value specified in EXT_CAPT_MASK parameter in the Condense configuration file, dtlca.cfg.

The Condense task creates the partial condense files with the following name:

```
hlq.CND.CPyymmdd.Thhmmssn
```

where:

```
hlq = EXT_CAPT_MASK value
yyymmdd = date
hhmmss = time
n = sequence number starting at 1 (for uniqueness)
```

Use one of the following modes to read condense files:

- ♦ batch change extraction mode with the CAPX access method
- ♦ continuous change extraction mode with the CAPXRT access method

For more information accessing condensed data and the extraction modes available for condense files, see “Accessing CDC Data” on page 55.

For information about configuring PowerExchange for continuous change extraction mode, see “Configuring PowerExchange for Oracle CDC” on page 10 .

You can use the PowerExchange Navigator and PowerCenter to extract the changes from condense files. For more information about using PowerCenter to extract changes, see *PowerExchange Interfaces for PowerCenter*.

Checkpoint Files

There must be at least two Checkpoint files used by the system with allocation parameters defined in the Condense configuration file, dtlca.cfg.

Each time the system starts up, a fresh Checkpoint records the results of merging the current registrations from the CCT file with the information from the latest checkpoint. If it is a Cold Start, then no Checkpoint files are brought forward.

After each file switch or shut down, a fresh checkpoint is taken.

The following table describes the information stored in Checkpoint files:

Checkpoint Record Type	Description
ERT records	Registration tags and restart tokens that indicate the point to start receiving records from the change stream.
DCT records	Information that is also held in the CDCT file, describing completed Condensed files. This information is purged based on the COND_CDCT_RET_P parameter in the Condense configuration file, dtlca.cfg. This record type enables you to restore the CDCT file to a point consistent with a restart.
SRT record	A single record defining system-wide information.

Configuring PowerExchange Condense Parameters

Table 6-1 describes the required parameters for the Condense configuration file, dtlca.cfg:

Table 6-1. Required Parameters

Parameter	Description	Valid Values
CAPT_IMAGE	Specifies whether before or after images are recorded.	AI = After images. BA = Both before and after images.
CHKPT_NUM	Specifies the number of checkpoint files. Warning: Decreasing CHKPT_NUM on a warm start can result in restart from an incorrect location. Only decrease CHKPT_NUM if doing a cold start. Default is 3.	1-999999
CHKPT_BASENAME	Base name for checkpoint files. To form the full checkpoint file name, Vn is appended, where <i>n</i> is 0 to CHKPT_NUM-1. If a fully-qualified path is not specified, the checkpoint files are created in the PowerExchange directory.	For example, INFA.D.CHKPT. Maximum length is 256.

Table 6-1. Required Parameters

Parameter	Description	Valid Values
COLL_END_LOG	<ul style="list-style-type: none"> - 0. Continuous mode. After each condense run, the system waits for the number of minutes defined in the NO_DATA_WAIT parameter, then performs another Condense. - 1. Batch mode. The system shuts down after a single condense run. For example, a single condense run might be scheduled following a particular batch update job. 	0 = Continuous 1 = Batch
CONDENSE_SHUTDOWN_TIMEOUT	Specifies the maximum time period for the PowerExchange Condenser, DTLCAON, to shut down normally after a shutdown command. This period should be set based on your environment, such as the number of tables being condensed.	0 to 2147483647 seconds. Default is 600 seconds.
COND_CDCT_RET_P	CDCT and Condensed files retention period in days. Files older than this period and their corresponding CDCT records are deleted during start-up, fileswitch, or shut down processing. Be sure to allow enough days for the files to be processed before they are deleted.	Any number > 0
CONN_OVR	Specifies the CAPI_CONNECTION name to use when running PowerExchange Condense. Otherwise, PowerExchange Condense uses the default connection.	Valid source CAPI connection name.
DBID	Specifies the PowerExchange registration collection identifier, also called instance name. Used in conjunction with DB_TYPE, it defines selection criteria for which registrations in the CCT file to select. It must match the instance name displayed for the Registration Group containing the capture registrations.	Collection identifier used for capture registrations
DB_TYPE	Database type.	ORA
EXT_CAPT_MASK	Unique mask for files created by the Condense. This should be specified in upper case. EXT_CAPT_MASK gives the condense process complete control over any condense files that match the specified EXT_CAPT_MASK. Warning: Do not share EXT_CAPT_MASK between multiple condense processes. Because each process considers all files its own, it might delete files associated with other condense processes. Do not re-use an EXT_CAPT_MASK until all files that match the mask have been finished with. Any file names that match the mask, even if unrelated to condense, are considered condense files by PowerExchange. Make sure your EXT_CAPT_MASK does not match any other files on the system.	For example: INF.A.D Maximum length is 256 characters

Table 6-1. Required Parameters

Parameter	Description	Valid Values
FILE_FLUSH_VAL	<p>Specifies the file flush interval in seconds. The file flush interval is the number of seconds that elapse before a flush is performed on the current partial condense file. When the Condense task flushes, the data is written to the disk condense files allowing it to be read by continuous extraction mode extractions.</p> <p>Valid values are:</p> <ul style="list-style-type: none"> -1 causes the Condense task to not flush data to the current file. Use this setting, or allow it to default, if using batch extraction mode for condense files. Change this setting if using continuous extraction mode for condense files. Otherwise, latency increases for continuous extraction mode extractions. This value is the default value. - 0 causes Condense to flush after every record. <p>Warning: Specifying 0 impacts the Condense task and file system performance.</p> <ul style="list-style-type: none"> - 1 to 86400 sets the file flush interval in seconds. 	-1 to 86400 seconds
FILE_SWITCH_CRIT	<p>Defines the criteria to use when deciding when to do an automatic file switch. Use in conjunction with FILE_SWITCH_VAL and FILE_SWITCH_MIN. Default is M.</p>	<p>M = Minutes R = Records</p>

Table 6-1. Required Parameters

Parameter	Description	Valid Values
FILE_SWITCH_MIN	<p>Specifies file switch criteria for Condense when changes for new sources are encountered. Use this to reduce latency for continuous extraction mode.</p> <p>Format is: (<i>min_val</i>,<i>min_val_ign</i>)</p> <p>Where: <i>min_val</i> specifies the minimum number of FILE_SWITCH_CRIT units that must pass before a file switch is done when encountering a change for a source with no entry in the CDCT.</p> <p>Valid values: -1 causes FILE_SWITCH_MIN to be ignored. File switch processing is then controlled by FILE_SWITCH_VAL only. - 0 causes the Condense task file to switch for every new source. - 1 to 2147483647 causes the Condense task file to switch every <i>n</i> units. <i>min_val_ign</i> specifies the number of FILE_SWITCH_CRIT units which must occur during cold start processing before Condense uses the <i>min_val</i> value. Until <i>min_val_ign</i> is exceeded, FILE_SWITCH_CRIT alone controls file switch activity. Valid values: - 0 causes the Condense task to use the minimum file switch value specified in <i>min_val</i> immediately on cold start. - 1 to 2147483647 causes the Condense task file to switch every <i>n</i> FILE_SWITCH_CRIT units. - The <i>min_val_ign</i> is not used when Condense is warm started. - Default is (-1,0). Warning: Specifying (0,0) results in large numbers of file switches during cold start, when many sources are registered in the CDCT.</p>	<p><i>min_val</i>: -1 to 2147483647. <i>min_val_ign</i>: 0 to 2147483647</p>
FILE_SWITCH_VAL	<p>Defines the number of FILE_SWITCH_CRIT units at which to perform a file switch automatically. Default is 30.</p> <p>For example, if this value is 30, the Condense task file switches every 30 records if FILE_SWITCH_CRIT=R or every 30 minutes if FILE_SWITCH_CRIT=M.</p> <p>If the condense file is empty of data at the FILE_SWITCH_VAL interval, then no file switch takes place.</p>	Any number > 0
GROUPDEFS	<p>Fully-qualified path and filename to the Condense Group Definition Parameter file containing condense group parameters. Use to enable the Flexible Condense capability.</p>	Maximum is 255 characters

Table 6-1. Required Parameters

Parameter	Description	Valid Values
NO_DATA_WAIT	<p>When running in Continuous mode, it defines the number of minutes to wait on commands manually entered through the Command Handler before starting the next Condense. Typically, the system might wait for 60 minutes before starting the next Condense.</p> <p>If file switching is performed based on time and FILE_SWITCH_VAL is smaller than NO_DATA_WAIT, then the wait period is reduced to the smaller of the two values.</p> <p>When the system runs in Batch mode (COLL_END_LOG = 1), this parameter is not used.</p> <p>Default is 60 minutes.</p>	Any number > 0
NO_DATA_WAIT2	<p>Defines the number of seconds before the Condenser stops.</p> <p>The completion of a Condense run occurs when this period expires without data being provided.</p> <p>The optimum value for setting the parameter varies according to the loading of the system.</p> <ul style="list-style-type: none"> - If the parameter is set too low, the Condense might report that no data when data exists. A delay might occur if a large unit of work is started, such as one containing several thousand rows. - If the parameter is set too high, then an excessive period of apparent inactivity elapses before control returns to the Command Handler, allowing you to enter commands such as FILESWITCH or SHUTDOWN. - 60 seconds is recommended. - Default is 600 seconds. 	Any number > 0
PWD	Password used to identify the user (see UID) and check their authorizations.	String
RESTART_TOKEN and SEQUENCE_TOKEN	<p>Every registration in the Change Data Extraction Repository file contains a restart token.</p> <p>The token has two parts: a sequence and a restart token.</p> <ul style="list-style-type: none"> - RESTART_TOKEN - SEQUENCE_TOKEN <p>The restart tokens can be provided at application level, registration level, or both. For more information, see "Creating Restart Points with DTLUAPPL" on page 35.</p>	
SIGNALLING	<ul style="list-style-type: none"> - SIGNALLING = Y. The system takes automatic action in the event of certain errors, such as memory corruption. The system attempts to close down in an orderly manner. - SIGNALLING = N. No automatic trapping of system errors is performed. The operating system uses default error handling, which is usually to report the offending program line and dump memory. <p>Default is N.</p>	<p>Y = Attempt to close down after error.</p> <p>N = Abort on error.</p>

Table 6-1. Required Parameters

Parameter	Description	Valid Values
UID	User ID to identify the user and check their authorizations. This User ID must be a valid Oracle User ID that has authority to access the Oracle redo logs and LogMiner.	String
VERBOSE	Specifies whether PowerExchange Condense issues verbose or terse messages for frequent condense activity like cleanup, checkpoints, condense cycles, and file switch processing. Valid values: - Y for verbose messaging - N for terse messaging - Default is Y.	Y or N

Configuring Flexible Condense

If you have a database containing customer information with three tables: Personnel, Departments, and Locations. You need all three tables for each of your 1,000 customers, resulting in 3000 tables.

To capture changes from these tables, a registration would have to be created for each schema version of the table, resulting in 3,000 registrations.

To reduce this registration creation work load, you can use Flexible Condense to create a single registration for a particular table type and then specify a Condense Group Definition file. In this file, you can specify the related groups, schemas and registrations that you want to associate with the registration. In our example above this would represent our 1,000 customers.

With Flexible Condense, for the previous example, you can create three registrations and one Condense Group Definition file to capture changes on all 3,000 tables.

Running Flexible Condense

Flexible Condense is configured by specifying the GROUPDEFS parameter in the Condense configuration file, dtlca.cfg. This identifies a sequential file which contains condense group definitions.

When a condense runs, the Condense Group Definition (CGD) file is processed and the resulting condensed files created.

Note: Any unregistered schema versions of the tables need to be registered with Oracle Log Miner supplemental log groups.

Condense Group Definition File

The Condense Group Definition (CGD) file is specified by the GROUPDEFS parameter in the Condense configuration file. It contains one or more GROUP statements, each including a SCHEMA and one or more REG keywords:

Keyword	Parameter	Type/Length	Description
GROUP	GroupName	VARCHAR(255)	Identification for the Condense group.
	External_capture_mask	VARCHAR(255)	Fully-qualified path prefix for the resulting condense files.

Keyword	Parameter	Type/Length	Description
SCHEMA	Schema name	VARCHAR(255)	Schema override.
REG	Registration	VARCHAR(8)	Full or wild-carded registration name. Must be the prefix.

When creating a Condense Group Definition file, note the following:

- ◆ Each specified GroupName must be unique.
- ◆ Each specified External_capture_mask must be unique.
- ◆ Multiple SCHEMA statements can be specified under each GROUP.
- ◆ A SCHEMA statement is not required for the schema used in the base registration, the first GROUP.
- ◆ REG statements apply to the preceding SCHEMA statement only. If no SCHEMA statement is present, the REG statements apply to the preceding GROUP statement.
- ◆ If a SCHEMA or REG statement is found without a preceding GROUP statement, a syntax error is generated.
- ◆ It is not possible to specify the same REG statement for multiple GROUP statements.
- ◆ If a REG statement is not specified for a group, then all the registrations belonging to that group are considered for condense. This is as if "REG=*" has been specified.

Example 1

An Oracle Instance (DBID) ORAINST has the following registrations:

Registration	Table Name
Reg_Emp1	COMPANY1.EMPLOYEES
Reg_Emp2	COMPANY1.EXEMPLOYEES
Reg_Manager	COMPANY1.MANAGERS
Reg_Loc1	COMPANY1.UK_LOCATIONS
Reg_Loc2	COMPANY1.US_LOCATIONS
Reg_Loc3	COMPANY1.JAPAN_LOCATIONS
Reg_Dept1	COMPANY1.DEPTS
Reg_Job1	COMPANY1.JOBS
Reg_Job2	COMPANY2.JOBS

Similar tables, COMPANY2.EMPLOYEES and COMPANY3.EMPLOYEES, exist for schemas COMPANY2 and COMPANY3. Registration Reg_Job2 is for COMPANY2.JOBS.

A simple solution is to group condense files by schema for all registrations. Note how a schema override is not required for the schema used in the base registration, Company1:

```
GROUP= (Company1, "/user/condense_files/company1")
GROUP= (Company2, "/user/condense_files/company2")
SCHEMA=Company2
GROUP= (Company3, "/user/condense_files/company3")
SCHEMA=Company3
```

The following input sequence employs an optional REG=(*) statement for clarity. It yields the same results:

```
GROUP= (Company1, "/user/condense_files/company1")
REG=*
GROUP= (Company2, "/user/condense_files/company2")
SCHEMA=Company2
REG=*
GROUP= (Company3, "/user/condense_files/company3")
```

```
SCHEMA=Company
REG=*
```

Note that Registration Reg_Job2 specifies source table COMPANY.JOBS. Using the SCHEMA override creates a duplicate. The changes are only condensed once. Also because GROUP Company1 and Company2 would contain changes for registration Reg_Job2, the changes are only be written to GROUP Company1.

Example 2

Note how the division of groups is more complex in this example:

```
GROUP= (Company1People, "/user/condense_files/people/company1")
REG=Emp*
REG=Manager
GROUP= (UK_People, "/user/condense_files/people/UK")
SCHEMA=Company2
REG=Manager
REG=Emp*
REG=Em* a
SCHEMA=Company3
REG=Manager
REG=Emp*
GROUP= (All_Managers, "/user/condense_files/people/managers")
SCHEMA=Company1
REG=Manager b
SCHEMA=Company2
REG=Manager
SCHEMA=Company3
REG=Manager
GROUP= (AllCompany3_Locations, "/user/condense_files/people/locations/company3")
REG=loc*
GROUP= (Company2Jobs, "/user/condense_files/Jobs/company2")
REG=Job* c
```

^a The order of registration statements is irrelevant. All registrations matching the wild cards for all REG statements will be grouped.

^b Condensed changes for the COMPANY2.MANAGER table are only written to the UK_people group and not to both UK_people and All_Managers groups.

^c The Company2Jobs group does not require a SCHEMA override because a base registration exists for table COMPANY2.JOBS. But this group will also contain the changes for COMPANY1.JOBS.

Output Files

The condense files are written to the paths specified by the EXTERNAL_CAPTURE_MASK parameter of the GROUP statement.

The condensed change data can then be accessed through an extraction map as normal.

Starting and Stopping Condense

Starting PowerExchange Condense

Use the following command to start PowerExchange Condense:

```
dtlcacon
[config=config_file]
[cs=cond_cfg_file]
```

```
[license=license_file]
```

Parameter	Description
<code>config_file</code>	Optionally specifies the DBMOVER configuration file name, including the full path. Default is <code>PWX_install_dir/dbmover.cfg</code> .
<code>cond_cfg_file</code>	Optionally specifies the PowerExchange Condense configuration file name, including the full path. Default is <code>PWX_install_dir/dtlca.cfg</code> .
<code>license_file</code>	Optionally specifies the license key file name, including the full path. Default is <code>PWX_install_dir/license.key</code> .

For more information about controlling PowerExchange Condense, see *PowerExchange Command Reference*.

The following rules and guidelines apply:

- ♦ To cold start remove all checkpoint files defined by the PowerExchange Condense configuration file parameter `CHKPT_BASENAME`.
- ♦ Cold start behaves differently depending upon the setting of the `RESTART_TOKEN` and `SEQUENCE_TOKEN` in the `dtlca.cfg` file:
 - If the `RESTART_TOKEN` and `SEQUENCE_TOKEN` are not present in the `dtlca.cfg` parameters, then PowerExchange Condense starts from the current position.
 - If the `RESTART_TOKEN` and `SEQUENCE_TOKEN` are set to zero in the `dtlca.cfg` parameters, then PowerExchange Condense starts from the most current Oracle catalog dump in the Oracle redo logs.
 - If the `RESTART_TOKEN` and `SEQUENCE_TOKEN` are to be set to a specific value, then PowerExchange Condense starts from the value specified.
- ♦ Create capture registrations using the PowerExchange Navigator for those sources for which change data capture is required.
- ♦ Edit `dtlca.cfg` to change the appropriate parameters. For more information, see “Configuring PowerExchange Condense Parameters” on page 46.
- ♦ Start PowerExchange Condense.

Shopping PowerExchange Condense

When running in foreground, use one of the following commands to stop PowerExchange Condense:

- ♦ **SHUTDOWN.** PowerExchange Condense performs a file switch, updates the CDCT, takes a checkpoint, and ends.
- ♦ **SHUTCOND.** PowerExchange Condense starts a final condense cycle. When the condense cycle ends, PowerExchange Condense performs a file switch, updates the CDCT, takes a checkpoint, and ends.

When running in background on Linux or UNIX, PowerExchange Condense processes the `KILL` command as `SHUTDOWN`.

For more information about controlling PowerExchange Condense, see *PowerExchange Command Reference*.

CHAPTER 7

Accessing CDC Data

This chapter includes the following topics:

- ♦ Overview, 55
- ♦ Oracle CDC Real-time Extraction Mode, 55
- ♦ Oracle CDC Condense File Extraction, 57

Overview

Extract captured changes from the Oracle redo logs by PowerExchange software. The capture processes available for accessing captured Oracle data are PowerExchange Change and the near real-time version, PowerExchange Real-time.

This chapter provides examples to help describe the process. For more information about the full command syntax, see the *PowerExchange Reference Manual*.

Oracle CDC Real-time Extraction Mode

PowerExchange Oracle CDC Real-time extraction mode is near real-time access to changes from the Oracle redo logs. Oracle CDC uses Oracle LogMiner to read the logs.

When registering a table through the PowerExchange Navigator for Real-time extraction mode only, the Condense option should be set to None.

At this point, the source tables should be part of a supplemental log group, LogMiner should be configured, and the tables registered for capture through the PowerExchange Navigator.

Testing with the Navigator

You can use the PowerExchange Navigator Row Test functionality to prove that changes are being captured from your registered tables.

To test if changes are being captured from registered tables:

1. Select the Extraction Group that holds definition that you require to test.
2. Select the extraction definition that you must test.

3. Click **File > Database Row Test** or click the **Row Test** button.
4. Enter the following information into the Row Test dialog box.

Field	Description
Location	The location for the target node as set in the dbmover.cfg file.
UserID/Password	The user ID and password used to access the data in Oracle
Application Name	The application that was created using DTLUAPPL. This is mandatory for CAPXRT. Note that this is case-sensitive.
SQL Statement	The SQL statement is automatically generated. The table is identified in the following format: d8Target Subsystem_id.captureRegistrationName_Tablename d8 always used for Oracle. For example: select * from d8oemcap.oradoc01_EMP
Target Subsystem id	From your capture registration.
CaptureRegistrationName	From your capture registration.
Tablename	From your capture registration.

5. Click **Go**.

The results of the data retrieval appear. These results represent the changes that have occurred on the table that you registered for data capture.

Oracle LogMiner logs a vast amount of data. If no data is returned when data is expected, adjusting the timeout parameter can allow the captured data to be returned.

6. To access this screen, click the **Advanced** button. For more information about the Advanced Parameters dialog box, see the *PowerExchange Navigator User Guide*.

Using Real-time Extraction Mode in PowerCenter

Using PowerExchange Client for PowerCenter, a workflow can extract Oracle changes in real-time extraction mode. Use PWX Oracle CDC Real-time application connections. For more information, see *PowerExchange Interfaces for PowerCenter*.

PowerExchange Listener Messages During Data Extract

An extract may take some time to complete. To help understand the progress, messages are directed to the PowerExchange Listener. A timeout parameter can be coded and shown in the Advanced Parameters dialog box. This timeout parameter is used as a timer at which progress will be reported to the listener on the extract, however, the extract will continue until the End Of Log is found. Sample messages are shown below:

```
PWX-09957 CAPI i/f: Read times out after 60 seconds
PWX-09966 CAPI i/f: No data in timeout period. Last data time=20030710115146
PWX-09966 CAPI i/f: No data in timeout period. Current time=03/07/10 16:52:20
PWX-09966 CAPI i/f: No data in timeout period. Current time=03/07/10 16:53:20
PWX-09966 CAPI i/f: No data in timeout period. Current time=03/07/10 16:54:21
PWX-09966 CAPI i/f: No data in timeout period. Current time=03/07/10 16:55:21
PWX-09966 CAPI i/f: No data in timeout period. Current time=03/07/10 16:56:21
PWX-09966 CAPI i/f: No data in timeout period. Current time=03/07/10 16:57:21
PWX-09966 CAPI i/f: No data in timeout period. Current time=03/07/10 16:58:21
PWX-09966 CAPI i/f: No data in timeout period. Current time=03/07/10 16:59:21
PWX-09967 CAPI i/f: End of log for time 03/07/10 16:49:22 reached
PWX-00408 DBAPI File d8oemdet.IfITestB_IFITESTB, 20 rows read, 7006 bytes input.
```

Two messages are output where CAPI supports Current End of Log processing - which will be all systems.

- ♦ 9967 is issued when the EOL is met, so is displayed once only.
- ♦ 9966 is issued when a CAPI_NODATA occurs before the EOL is met. If any records were read, then a data time stamp is displayed otherwise the current time is displayed.

Oracle CDC Condense File Extraction

PowerExchange Oracle CDC accesses changes from condense files created by the PowerExchange Condense task. You can access the change data in batch extraction mode using the CAPX access method. You can also access the change data in continuous extraction mode using the CAPXRT access method.

The PowerExchange Condense tasks creates the condense files by reading the changes from Oracle redo logs using Oracle LogMiner. You specify the frequency that the data is condensed using parameters in the Condense task configuration file. See “Configuring PowerExchange Condense Parameters” on page 46.

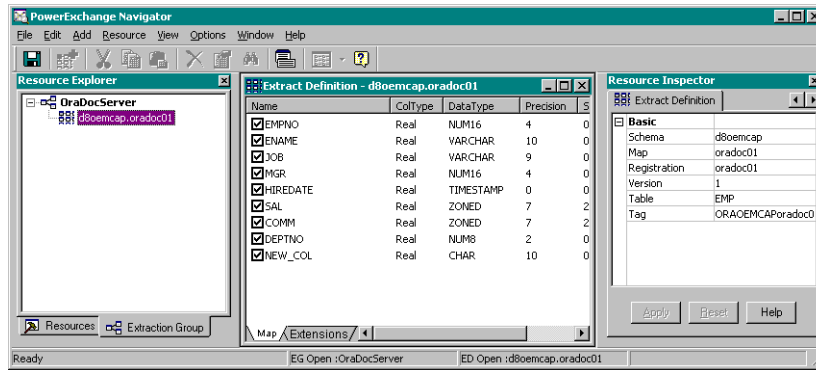
In order for a capture registration to be eligible for Condense, the Condense option must be set to Part. For more information about modifying existing registrations to take advantage of the Condense option, see “Amending CDC Definitions” on page 39.

Testing with the Navigator

Once a condense process is complete, you can see the data through the PowerExchange Navigator. You can use the Row Test functionality to prove that the changes are being captured from your registered tables.

To test for captured changes:

1. Select the Extraction Group that holds definition that you want to test
2. Select the extraction definition that you want to test.



3. Click **File > Database Row Test**.
4. Enter the following information into the Row Test dialog box.

Field	Description
Location	The location for the target node as set in the dbmover.cfg file.
UserID/Password	User ID and password used to access the data
Application Name	This is mandatory for CAPX and CAPXRT. For testing purposes this can be set to any string.
SQL Statement	The SQL statement is automatically generated. The table is identified in the following format: d8Target Subsystem_id.CaptureRegistrationName_Tablename d8 always used for Oracle. For example: select * from d8oemcap.oradoc01_EMP

5. To test batch extraction mode, select DB Type CAPX. To test continuous extraction mode, select DB Type CAPXRT.
6. Click **Advanced**.

If testing continuous extraction mode, specify the CAPX CAPI_CONNECTION name in CAPI Connection Name.

By default, PowerExchange does not treat the row test like a true update. If you want the test to be treated as a true extract and you want to update the restart information under an application, clear the No Progress Update option. When this extract runs, an entry appears under the Application group for the application entered in the screen above.

Note: If you clear the No Progress Update option, the returned data will not be available for a “Since Last” data extract unless the start point is reset. Use this option with caution. If the progress indicator has been cleared by mistake, you can reset the start point of the extract through the Navigator. For more information, see “Reset to New Start Point” on page 38.

For more information about the Advanced Parameters dialog box, see the *PowerExchange Navigator User Guide*.

7. Click **Go**.

The results of the data retrieval are displayed. These represent the changes that have occurred on the table that you registered for data capture.

Using Batch Extraction Mode in PowerCenter

Using PowerExchange Client for PowerCenter, a workflow can extract Oracle changes in batch extraction mode. Use PWX Oracle CDC Change application connections. For more information, see *PowerExchange Interfaces for PowerCenter*.

Using Continuous Extraction Mode in PowerCenter

Using PowerExchange Client for PowerCenter, a workflow can extract Oracle changes in continuous extraction mode. Use PWX Oracle CDC Real Time application connections and specify the name of the CAPX CAPI_CONNECTION in the CAPI Connection Name Override attribute. For more information, see the *PowerExchange Interfaces for PowerCenter*.

CHAPTER 8

Stopping Change Capture

This chapter includes the following topic:

- ♦ Stopping Change Capture, 61

Stopping Change Capture

Capture of Oracle data is through the Supplemental Log Group and Oracle redo logs. PowerExchange can then extract that information and either apply it directly to a target database or make it available to partner software to manipulate and load. You can stop the collection of data with the following alter statement:

```
ALTER TABLE table DROP SUPPLEMENTAL LOG GROUP
```

Note: When you stop data collection, changes to data are not recorded. If the supplemental log group is to be reinstated then rematerializing the target database may be the only sensible choice.

INDEX

A

- access methods
 - CAPX 1
 - CAPXRT 1
- ALTER SYSTEM SET statement
 - setting the Oracle transaction_auditing parameter 8
- application groups
 - defined 38
- applications
 - adding DTLUAPPL configuration information 36
 - changing DTLUAPPL configuration information 37
 - printing status with DTLUAPPL 37
- Archivelog mode
 - Oracle requirement 8

B

- baseline
 - creating for data capture with DTLUAPPL utility 35
- Batch mode
 - for condensing changes 44

C

- CAPICAPX parameters 19
ORCL parameters 15
statements 14
- CAPICAPI configuration parameters
See PowerExchange Reference
- CAPL_CONN_NAME
CAPICAPI statement 14
- CAPL_CONNECTION
CAPICAPI statement 14
- capture
 - required authority 9
- capture definitions
 - updating status or comments 40
- capture registrations
 - adding 28
 - viewing 29
- capture_connect_string
ORACLEID parameter 13
- capturing data
 - DTLUAPPL utility 35
 - registering 46
- CAPX
 - access method 1

- starting 53
- CAPX parameters
 - CAPICAPI 19
- CAPXRT
 - access method 1
- CATBEGIN
 - parameter 17
- CATEND
 - parameter 17
- CATINT
 - parameter 17
- CDC
 - phases 2
- CDCT file
 - condense output file 45
- change capture
 - registering condense 43
 - stopping 61
 - with PowerExchange Change and PowerExchange Real-time 1
- changed data capture
 - phases 2
- changing
 - data sources and registrations 40
- Checkpoint Datasets 46
- Checkpoint files
 - condense output files 46
 - record types 46
- CHKPT_BASENAME
 - parameter 54
- client binaries
 - Oracle requirement 8
- cold start
 - effect on Checkpoint files 46
- collection_id
 - ORACLEID parameter 13
- columns
 - removing from extraction maps 39
- command line
 - running DTLUAPPL with UNIX 35
 - running DTLUAPPL with Windows 35
- condense
 - configuration file parameters 46
 - flexible 51
 - migration 44
 - output files 45
 - registering 43
 - shutting down 54
- Condense files
 - condense output files 45

- Condense Group Definition (CGD) file
 - Flexible Condense 51
- Condense Group Definition (CGD) files
 - Flexible Condense processing 51
- condense output files
 - Checkpoint files 46
 - Condense files 45
- condensing changes
 - operational modes 44
- configuration file
 - condense parameters 46
- configuration files
 - DTLUAPPL utility 35
- configuring
 - Flexible Condense 51
- connect_string
 - ORACLEID parameter 14
- Continuous mode
 - for condensing changes 44

D

- data
 - steps for sourcing Oracle 5
- data delivery
 - format 34
- data extract
 - Listener messages 57
- data sources
 - updating capture registrations to reflect changes 40
- DCT records
 - in Checkpoint files 46
 - matching CDCT records 45
- deleting
 - registration entries 29
 - registration groups 28
- disabling
 - capture registrations 40
- dtlca.cfg
 - configuring Flexible 51
- DTLDSN
 - extraction criteria 32
- dtltknp.txt
 - adding application information 36
 - changing application information 37
 - sample configuration file 35
- DTLUAPPL
 - examples 36
- DTLUAPPL utility
 - adding application information 36
 - capturing data 35
 - changing application information 37
 - printing application status information 37
 - running against remote databases 37
 - using with UNIX 35
 - using with Windows 35

E

- ERT records
 - in Checkpoint files 46

- extraction criteria
 - specifying 32
- extraction definition
 - columns 33
- extraction maps
 - updating to reflect data source changes 39

F

- FILE_SWITCH_CRIT
 - file switch tag 45
- FILE_SWITCH_VAL
 - file switch tag 45
- FILESWITCH 44
- Flexible Condense
 - Condense Group Definition (CGD) file 51
 - configuring for 51
 - examples 52
 - testing rows 57
 - using 51
- formats
 - data delivery 34

G

- GROUPDEFS
 - Flexible Condense parameter 51

H

- handling unformatted data
 - BYPASSUF parameter 15

L

- Listeners
 - data extract messages 57
 - testing remote 5
- LOB columns
 - BYPASSUF parameter 15
- log files
 - DTLUAPPL utility 35
- LogMiner
 - configuration requirement 10
 - environment for sourcing data 6

M

- migration
 - condense 44
- modes
 - operational for condensing changes 44

N

NO_DATA_WAIT 44

O

ODBC
 interface to PowerExchange 32
Oracle
 steps for sourcing data 5
Oracle catalog
 copying 10
 copying to the Redo log 8
Oracle SQL *Loader
 restrictions 23
Oracle_SID
 ORACLEID parameter 14
ORACOLL
 DTLUAPPL parameter 36
ORACON
 DTLUAPPL parameter 36
ORAINST
 DTLUAPPL parameter 36
ORASchema
 DTLUAPPL parameter 36
ORCL parameters
 CAPI 15
output files
 CDCT file 45
 condense 45

P

parameters
 condense configuration 46
password
 personal metadata 6
personal metadata
 creating 6
PowerExchange Batch
 overview 1
PowerExchange Change
 change capture 1
PowerExchange Real-time
 change capture 1
printing
 application status information with DTLUAPPL 37

R

Real Application Cluster (RAC) environment
 CAPI ORCL parameter 16
record types
 in Checkpoint files 46
Redo log
 Oracle catalog 8
registering
 captured data 46
registration
 migration 44

registration entries
 deleting 29
registration groups
 adding 26
 deleting 28
 viewing 29
registration process
 steps 25
required authority
 for capture 9
requirements
 active LogMiner environment 10
 copying Oracle catalog to Redo log 8
 copying the Oracle catalog 10
 for sourcing Oracle data 5
 Oracle Archivelog mode 8
 Oracle client binaries 8
 Oracle supplemental logging 8, 9
 Oracle transaction_auditing parameter 8
resetting
 start points 38
RESTART_TOKEN
 parameter 54
restrictions
 SQL *Loader 23
row testing
 captured changes 55
 for Flexible Condense 57
RS
 extraction type 32

S

schemas
 managing changes 40
SEQUENCE_TOKEN
 parameter 54
shutting down
 condense 54
SL
 extraction type 32
SQL *Loader
 restrictions 23
SRT record
 in Checkpoint files 46
start points
 resetting 38
starting
 CAPX 53
status
 capture definitions 40
stopping
 change capture 61
supplemental logging
 Oracle requirement 8, 9

T

testing
 change capture 55
 remote Listeners 5

transaction_auditing parameter
Oracle requirement 8

U

UNIX
 running DTLUAPPL 35
updating
 extraction maps to reflect data source changes 39

W

Windows
 running DTLUAPPL 35

NOTICES

This Informatica product (the "Software") includes certain drivers (the "DataDirect Drivers") from DataDirect Technologies, an operating company of Progress Software Corporation ("DataDirect") which are subject to the following terms and conditions:

1. THE DATADIRECT DRIVERS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.

2. IN NO EVENT WILL DATADIRECT OR ITS THIRD PARTY SUPPLIERS BE LIABLE TO THE END-USER CUSTOMER FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR OTHER DAMAGES ARISING OUT OF THE USE OF THE ODBC DRIVERS, WHETHER OR NOT INFORMED OF THE POSSIBILITIES OF DAMAGES IN ADVANCE. THESE LIMITATIONS APPLY TO ALL CAUSES OF ACTION, INCLUDING, WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS.

